(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 February 2001 (01.02.2001)

PCT

(10) International Publication Number
WO 01/07476 A1

(51) International Patent Classification[7]: C07K 14/47, C12N 5/10, 5/16, 15/12, 15/63, 15/64

(21) International Application Number: PCT/US00/19666

(22) International Filing Date: 20 July 2000 (20.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/144,972	21 July 1999 (21.07.1999)	US
60/148,681	13 August 1999 (13.08.1999)	US
60/149,173	17 August 1999 (17.08.1999)	US
60/158,004	6 October 1999 (06.10.1999)	US
60/194,689	5 April 2000 (05.04.2000)	US

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(71) Applicant and
(72) Inventor: ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: BIRSE, Charles [GB/US]; 13822 Saddleview Drive, North Potomac, MD 20878 (US).

(74) Agents: WALES, Michele, M. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/07476 A1

(54) Title: 26 HUMAN PROSTATE AND PROSTATE CANCER ASSOCIATED PROTEINS

(57) Abstract: The present invention relates to newly identified human prostate cancer related polynucleotides and the polypeptides encoded by these polynucleotides (prostate cancer antigens). Also provided are vectors, host cells, antibodies, and recombinant methods for producing human prostate cancer antigens. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human prostate cancer antigens.

26 Human Prostate and Prostate Cancer Associated Proteins

Field of the Invention

This invention relates to newly identified prostate and prostate cancer related polynucleotides and the polypeptides encoded by these polynucleotides herein collectively known as "prostate cancer antigens", and the use of such prostate antigens for detecting disorders of the reproductive system, including disorders of the prostate, particularly the presence of cancer. This invention relates to prostate cancer antigens as well as vectors, host cells, antibodies directed to prostate cancer antigens and recombinant and synthetic methods for producing the same. Also provided are diagnostic methods for diagnosing and treating, preventing and/or prognosing disorders related to the prostate, including cancer, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of prostate cancer antigens of the invention. The present invention further relates to methods and/or compositions for inhibiting the production and/or function of the polypeptides of the present invention.

Background of the Invention

Cell growth is a carefully regulated process which responds to specific needs of the body. Occassionally, the intricate, and highly regulated controls dictating the rules for cellular division break down. When this occurs, the cell begins to grow and divide independently of its homeostatic regulation resulting in a condition commonly referred to as cancer. In fact, cancer is the second leading cause of death among Americans aged 25-44.

Prostate cancer has become the most common cancer among American men, and only lung cancer is responsible for more cancer deaths (Boring, Cancer Statistics, 41:19-36 (1991)). The age specific mortality rate has slowly increased over the past 50 years and in black American men is nearly double the rate found in white men (Carter, Prostate, 16:39-48 (1990)). Prostate cancer is responsible for nearly three percent of all deaths in men over the age of 55

years (Seidman, et al., Probabilities of Eventually Developing or Dying of Cancer-United States, 35:36-56 (1985)). Since the incidence of prostate cancer increases more rapidly with age than any other cancer, and the average age of American men is rising, the number of patients with prostate cancer is expected to increase dramatically over the next decade.

Approximately 30% of men with prostate cancer have distant metastases at the time of diagnosis (Schmidt, et al., J. Urol., 136:416-421 (1986)). Despite the impressive symptomatic response of metastases to hormonal manipulation (androgen deprivation), the survival rate for these patients is dismal: the median duration of survival is less than three years (Eyar, Urologic Pathology: The Prostate, Philadelphia, Pa., Lea and Febiger, 241-267 (1977)). By five years, over 75% and by ten years, more than 90% of these patients die of their cancer rather than with it (Silverberg, Cancer, 60:692-717 (1987) (Suppl.)). The problem with prostate cancer is that many forms of prostate cancer are latent, in other words, such forms are difficult to detect. Approximately 30% of the men over the age of 50 years who have no clinical evidence of prostate cancer harbor foci of cancer within the prostate (McNeal, et al., The Lancet, January, 11:60-63 (1986)). This remarkably high prevalence of prostate cancer at autopsy, seen in no other organ, makes it the most common malignancy in human beings (Dhom, J. Cancer Res. Clin. Oncol., 106:210-218 (1983)). There is strong support for the concept of multi-step process in the pathogenesis of prostate cancer in which latent cancers progress through some but not all of the steps necessary for full malignant expression (Utter, et al., J. Urol., 143:742-746 (1990).

There are a variety of techniques for early detection and characteristics of prostate cancers, however, none of them are devoid of problems. Prostate cancer is a notoriously silent disease with few early symptoms. There is a need, therefore, for identification and characterization of factors that modulate activation and differentiation of prostate cells, both normally and in disease states. In particular, there is a need to isolate and characterize additional molecules that mediate apoptosis, DNA repair, tumor-mediated angiogenesis, genetic imprinting, immune responses to tumors and tumor antigens and, among other things, that can play a role in detecting, preventing, ameliorating or correcting dysfunctions or diseases related to the prostate.

Summary of the Invention

The present invention includes isolated nucleic acid molecules comprising, or alternatively, consisting of, a prostate and prostate cancer associated polynucleotide sequence disclosed in Table 1 and the Sequence Listing and/or contained in a human cDNA clone described in Table 1 and deposited with the American Type Culture Collection ("ATCC"). Fragments, variants, and derivatives of these nucleic acid molecules are also encompassed by the invention. The present invention also includes isolated nucleic acid molecules comprising, or alternatively consisting of, a polynucleotide encoding a prostate and prostate cancer associated polypeptide. The present invention further includes prostate and prostate cancer polypetides encoded by these polynucleotides. Further provided for are amino acid sequences comprising, or alternatively, consisting of, prostate and prostate cancer associated polypeptides as disclosed in the sequence listing and/or encoded by the human cDNA clones described in Table 1 and deposited with the ATCC. Antibodies that bind these polypeptides are also encompassed by the invention. Polypeptide fragments, variants, and derivatives of these amino acid sequences are also encompassed by the invention, as are polynucleotides encoding these polypeptides and antibodies that bind these polypeptides. Also provided are diagnostic methods for diagnosing and treating, preventing and/or prognosing disorders related to the prostate, including cancer, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of prostate cancer antigens of the invention.

Detailed Description

Tables

Table 1 summarizes ATCC Deposits, Deposit dates, and ATCC designation numbers of deposits made with the ATCC in connection with the present application.

Table 2 indicates public ESTs, of which at least one, two, three, four, five, ten, or more of any one or more of these public EST sequences are optionally excluded from certain embodiments of the invention.

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide. The term "isolated" does not refer to genomic or cDNA libraries, whole cell total or mRNA preparations, genomic DNA preparations (including those separated by electrophoresis and transferred onto blots), sheared whole cell genomic DNA preparations or other compositions where the art demonstrates no distinguishing features of the polynucleotide/sequences of the present invention.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC, herein refered to as ATCC Deposit No.:Z. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having an amino acid sequence encoded by a polynucleotide of the invention as broadly defined (obviously excluding poly-Phenylalanine or poly-Lysine peptide sequences which result from translation of a polyA tail of a sequence corresponding to a cDNA).

In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000 continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain

coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number (e.g., ATCC Deposit No.: PTA-416 deposited July 23, 1999; ATCC Deposit No.: PTA-497 deposited August 11, 1999; ATCC Deposit No.: PTA-498 deposited August 11, 1999; and ATCC Deposit No.: PTA-1620 deposited April 3, 2000). The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, or the complement thereof (e.g., the complement of any one, two, three, four, or more of the polynucleotide fragments described herein), and/or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42 degree C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65 degree C.

Also included within "polynucleotides" of the present invention are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For

example, lower stringency conditions include an overnight incubation at 37 degree C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50 degree C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using oligo dT as a primer).

The polynucleotides of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and

RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000 continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

The polypeptides of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation,

formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1. "ATCC Deposit No:" refers to the ATCC Deposit which contains a particular cDNA Clone ID specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention).

Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:

RPTRPPCHILLAYLFFLWLCMAFLQVLSRYHSANHCYRMVTSFVLTVQQQIWVRLNLS VNFFFWCFFGLMTVSL (SEQ ID NO: 109);
FLQVLSRYHSANHCYRMVTSFVLTVQQQIWVRLNLSVNFFFWCFFGLMTVSLLYPCFA CNDSCMVFLTS (SEQ ID NO: 113); and/or
LVIYSWHXFFSFGFAWLFLQVLSRYHSANHCYRMVTSFVLTVQQQIWVRLNLSVNFFF WCFFGLMTVSL (SEQ ID NO:85). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

Included in this invention as preferred domains are Tachykinin family signature domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). Tachykinins, like most other active peptides, are synthesized as larger protein precursors that are enzymatically converted to their mature forms. Tachykinins are from ten to twelve residues long. A signature pattern, consisting of the last five residues of the C-terminus, which are conserved and are essential to the biological activity was used as the consensus pattern. The consensus pattern is as follows: F-[IVFY]-G-[LM]-M-[G>]. If the sequence is processed, the peptide ends with a C-terminal amidated Met while in a precursor sequence it is always followed by a Gly which subsequently provides the amide group. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: WCFFGLM (SEQ ID NO: 110); FFGLMTV (SEQ ID NO: 111); and/or WCFFGLMTV (SEQ ID NO: 112). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also

encompassed by the invention. Further preferred are polypeptides comprising one member of the Tachykinin family signature domains listed above, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of the amino acid sequence referenced in Table 1 for this gene. The additional contiguous amino acid residues may be N-terminal or C-terminal to the Tachykinin family signature domain. Alternatively, the additional contiguous amino acid residues may be both N-terminal and C-terminal to the Tachykinin family signature domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is characteristic of a signature specific to Tachykinin family members. Tachykinins are a group of biologically active peptides which excite neurons, evoke behavioral responses, are potent vasodilatators and secretagogues and contract (directly or indirectly) many smooth muscles. Based on the predicted tachykinin family signature domain, the translation product of this clone is expected to share at least some biological activities with tachykinins. Such activities are known in the art, some of which are described elsewhere herein. Tachykinin biological activity is easily assayed by techniques known in the art such as, e.g., Comis, et al., J Pept Res. 53(3):337-42 (1999).

The gene encoding the disclosed cDNA is believed to reside on chromosome 5. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 5.

This gene is expressed primarily in prostate and prostate cancer tissue as determined by expression analysis described in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the male reproductive tract, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen,

lymph, vaginal pool, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of the immunogenic epitopes shown in SEQ ID NO: 60 as residues: Tyr-10 to Cys-16. Polynucleotides encoding said polypeptides are encompassed by the invention.

It is known that the innervation of the prostate gland and seminal vesicle by various neuroactive peptides may be involved in the autonomic regulation of these organs in adult man, as well as sympathetic and parasympathetic nerve fibers (See, e.g. Lange W, et al., Urol Res. 18:337-40 (1990); which is hereby incorporated by reference herein). Additionally, studies by Sastry BV, et al. (Ann N Y Acad Sci. 632:339-53 (1991); which is hereby incorporated by reference herein) suggest that substance P-like tachykinins may play a role in sperm maturation, in expulsion of fluid from the epididymis, and in initiation of motility. Based on the tissue distribution in prostate and the presense of the predicted tachykinin signature domain, polypeptides of the invention are expected to share at least some biological activities with tachykinin family members and may be involved in normal prostate function. Therefore, polynucleotides and polypeptides corresponding to this gene would be useful for diagnosis, detection, treatment and/or prevention of disorders of the male reproductive system, including but not limited to protsate cancer. Additionally, the tissue distribution in prostate cancer tissue, indicates that polynucleotides and polypeptides of the invention would be useful for the treatment and diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate polynucleotides and polypeptides of the invention could be used to treat and diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to

its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1186 of SEQ ID NO:11, b is an integer of 15 to 1200, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of this clone is a cation dependent mannose 6 phosphate (CD-M6P) specific receptor, known to be involved in transport of lysosomal enzymes to lysosomes (See, e.g., Sohar I, et al., Biochem J. 330(Pt 2):903-8 (1998) and Ma, et al., J. Biol. Chem. 266, 10589-10595 (1991); which are hereby incorporated by reference herein).

The present invention relates to the novel finding of CD-M6P expression in prostate and prostate cancer tissues. Preferred embodiments of the invention are polynucleotide and polypeptide fragments, derivatives, and/or variants (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists, and antagonists of CD-M6P (such as, for example, antibodies directed to CD-M6P or fragments thereof) which would be useful for the treatment, prevention, detection and/or diagnosis of disorders of the prostate, including prostate cancer.

In specific embodiments, a preferred polypeptide fragment is the processed extracellular domain of CD-M6P, comprising, or alternatively, consisting of the following amino acid sequence:
SWQTEEKTCDLVGEKGKESEKELALVKRLKPLFNKSFESTVGQGSDTYIYIFRVCREAG
NHTSGAGLVQINKSNGKETVVGRLNETHIFNGSNWIMLIYKGGDEYDNHCGKEQRRA
VVMISCNRHTLADNFNPVSEERGKVQDCFYLFEMDSSLACSPEISHLSVGS (SEQ ID NO: 114). In additional specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: SWQTEEKTCDLVGEKGKESEKELALVKRLKPL (SEQ ID NO: 115), FNKSFESTVGQGSDTYIYIFRVCREAGNHTSG (SEQ ID NO: 116), AGLVQINKSNGKETVVGRLNETHIFNGSNWIM (SEQ ID NO: 117), LIYKGGDEYDNHCGKEQRRAVVMISCNRHTLA (SEQ ID NO: 118), and/or DNFNPVSEERGKVQDCFYLFEMDSSLACSPEISHLSVGS (SEQ ID NO: 119). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer tissues as determined by expression analysis described in Example 3. Additionally, this gene is expressed in dendritic cells.

Polynucleotide and polypeptide fragments, derivatives, and/or variants (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists, and antagonists of CD-M6P (such as, for example, antibodies directed to CD-M6P or fragments thereof) of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for

diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, and immune system disorders. Similarly, polypeptide of the invention and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, immune, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, immune, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, or all four of the immunogenic epitopes shown in SEQ ID NO: 61 as residues: Glu-23 to Thr-31, Glu-37 to Lys-44, Asn-94 to Lys-99, Phe-112 to Phe-117. Polynucleotides encoding said polypeptides are encompassed by the invention.

The finding of CD-M6P expression in prostate and prostate cancer tissue, indicates that polynucleotide and polypeptide fragments corresponding to this gene, derivatives, and/or variants thereof (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists, and antagonists of CD-M6P (such as, for example, antibodies directed to CD-M6P or fragments thereof) would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer. The expression in the prostate tissue may indicate that polynucleotide and polypeptide fragments corresponding to this gene, derivatives, and/or variants thereof, agonists, and antagonists of CD-M6P could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, polynucleotide

and polypeptide fragments, agonists and antagonists may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein fragments, agonists and antagonists, as well as, antibodies directed against the same may show utility as a tumor marker and/or immunotherapy targets for the prostate.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1092 of SEQ ID NO:12, b is an integer of 15 to 1106, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of this gene shares sequence homology with human Zn-alpha 2-glycoprotein (ZAG); and may represent a prostate specific variant of ZAG (See, e.g., Genbank Accession No. gi|467671|emb|CAA49574.1|; all references available through this accession are hereby incorporated by reference herein). ZAG stimulates lipid degradation in adipocytes and causes the extensive fat losses associated with some advanced cancers. ZAG shares 30-40% sequence identity with the heavy chains of class I major histocompatibility complex (MHC) proteins. Therefore, included in this invention as preferred domains are immunoglobulins and major histocompatibility complex proteins signature domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). It is known that the Ig constant chain domains and a single extracellular domain in each type of MHC chains are related. These homologous domains are approximately one hundred amino acids long and include a conserved intradomain disulfide bond. A small pattern around the C-terminal cysteine involved in this

disulfide bond was developed which can be used to detect these category of Ig related proteins. The consensus pattern is as follows: [FY]-x-C-x-[VA]-x-H. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: YSCHVQH (SEQ ID NO: 120), PYSCHVQHS (SEQ ID NO: 121), and/or APYSCHVQHSS (SEQ ID NO: 122). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. Further preferred are polypeptides comprising the Immunoglobulins and major histocompatibility complex proteins signature domain listed above, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of the amino acid sequence reference in Table I for this gene. The additional contiguous amino acid residues may be N-terminal or C- terminal to the Immunoglobulins and major histocompatibility complex proteins signature domain. Alternatively, the additional contiguous amino acid residues may be both N-terminal and C-terminal to the Immunoglobulins and major histocompatibility complex proteins signature domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is a signature specific to Immunoglobulins and major histocompatibility complex related proteins.

The present invention relates to the novel finding of expression of a ZAG homolog in prostate and prostate cancer tissue. Preferred embodiments of the invention are polynucleotide and polypeptide fragments corresponding to this gene; variants and/or derivatives of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides); and agonists and/or antagonists of the polypeptides of the invention (including antibodies directed to the polypeptides of the invention

or to fragments thereof) which would be useful in the treatment, prevention, detection and/or diagnosis of disorders of the prostate, including prostate cancer. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:

PVLLSLLLLLGPAVPQENQDGRYSLTYIYTGLSKHVEDVPAFQALXHSMTSSSLDTTVK TGSLSPWDSGD RWKEWRIGSRTANFRRPGRTSLWRP (SEQ ID NO:86). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate or prostate cancer tissue as determined by expression analysis described in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the male reproductive system, such as prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, or all six of the immunogenic epitopes shown in SEQ ID NO: 62 as residues: Val-19 to Tyr-28, Tyr-61 to Gln-69, Met-80 to Arg-93, Tyr-107 to Gly-113, Glu-126 to Ser-

131, Tyr-138 to Asp-143. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and prostate cancer tissue, and homology to ZAG, indicates that polynucleotides and polypeptides corresponding to this gene, variants and/or derivatives of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides); and/or agonists and/or antagonists of these polypeptides (including antibodies directed to polypeptides or polynucleotides of the invention or to fragments thereof) would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression of this novel ZAG variant in the prostate tissue may indicate the gene or its products could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 873 of SEQ ID

NO:13, b is an integer of 15 to 887, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The gene encoding the disclosed cDNA is believed to reside on chromosome 17 at interval D17S922-D17S798 . Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 17.

The polypeptide encoded by gene has been determined to have transmembrane domains at amino acid residues from about 38 to about 54, from about 56 to about 72, from about 131 to about 147, and from about 154 to about 170 of the amino acid sequence referenced in Table 1 for this gene. Based upon these characteristics, it is believed that the protein product of this gene shares structural features to type IIIa membrane proteins. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:

MLQQDSNDDTEDVSLFDAEEETTNRPRKAKIRHPVASFFHLFFRVSAIIVYLLCELLSSS FITCMVTIILLLSCDFWAVKNVTGRLMVGLRWWNHIDEDGKSHWVFESRKESSQENK TVSEAESRIFWLGLIACPVLWVIFAFSALFSFRVKWLAVVIMGVVLQGANLYGTSGVR CAAESI (SEQ ID NO:87) and/or

SCDFWAVKNVTGRLMVGLRWWNHIDEDGKSHWVFESRKESSQENKTVSEAESRIFW LG (SEQ ID NO: 123). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate or prostate cancer tissue as determined by expression analysis described in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, especially disorders of the prostate, including but not limited to, prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, or all four of the immunogenic epitopes shown in SEQ ID NO: 63 as residues: Leu-2 to Glu-11, Ala-18 to Lys-30, Glu-107 to Val-119, Phe-194 to Ser-205. Polynucleotides encoding said polypeptides are also encompassed by the invention

The tissue distribution in prostate and/or prostate cancer tissues indicates that polynucleotide and polypeptides corresponding to this gene, fragments, variants or derivatives thereof (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists of this gene or gene products (including antibodies directed to polynucleotides or polypeptides of the invention or fragments thereof) would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate polynucleotide and polypeptides, fragments thereof, agonists and/or antagonists of the polypeptide encoded by this gene (including antibodies directed to polypeptides of the invention or fragments thereof) could be used to treat, prevent, detect, and/or diagnose disorders of the prostate, including

inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions.

The present invention relates to the novel finding of the expression of this gene in prostate and prostate cancer tissue. Preferred embodiments of the invention are polynucleotide and polypeptide fragments, variants or derivatives thereof (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists of this gene or gene products (including antibodies directed to the polynucleotides or polypeptides of the invention or to fragments thereof) which could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including prostate cancer. Moreover, the expression of this gene in prostate and prostate cancer tissues indicates that polynucleotides and/or polypeptides of the invention would be a good target for antagonists, particularly small molecules or antibodies, which inhibit its normal function. Accordingly, preferred are antibodies and or small molecules which specifically bind an extracellular portion of the translation product of this gene. The extracellular regions can be ascertained from the information regarding the transmembrane domains as set out above. Also provided is a kit for detecting prostate cancer. Such a kit comprises in one embodiment an antibody specific for the translation product of this gene bound to a solid support. Also provided is a method of detecting prostate cancer in an individual which comprises a step of contacting an antibody specific for the translation product of this gene to a bodily fluid from the individual, for example, serum or semen, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are more particularly described elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a

nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1904 of SEQ ID NO:14, b is an integer of 15 to 1918, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

This gene encodes the Human RAB-2 protein (See, Geneseq Accession No. W88104), a small GTP-binding protein which have been implicated in the regulation of many dynamic cellular processes (See, Goldenring et al., J. Biol. Chem. 268, 18419-18422 (1993)).

The present invention relates to the novel finding of HRAB-2 expression in prostate and prostate cancer tissue. Preferred embodiments of the invention are polynucleotide and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies to the polynucleotides or polypeptides corresponding to this gene) which can be used to treat, prevent, detect and/or diagnose disorders of the prostate, including prostate cancer.

Included in this invention as preferred domains are GTP-binding domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). From sequence comparisons and crystallographic data analysis it has been shown that an appreciable

proportion of proteins that bind ATP or GTP share a number of more or less conserved sequence motifs. The best conserved of these motifs is a glycine-rich region, which typically forms a flexible loop between a beta-strand and an alpha-helix. This loop interacts with one of the phosphate groups of the nucleotide. This sequence motif is generally referred to as the 'A' consensus sequence or the 'P-loop'. The consensus pattern derived from this motif is as follows: [AG].{4}GK[ST]. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: GESGVGKT (SEQ ID NO: 124), IGESGVGKT (SEQ ID NO: 125), GESGVGKTN (SEQ ID NO: 126), and/or VVLIGESGVGKTNLL (SEQ ID NO: 127). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. Further preferred are polypeptide fragments comprising the GTP-binding domains listed above, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of HRAB-2. The additional contiguous amino acid residues may be N-terminal or C- terminal to the GTP-binding domain. Alternatively, the additional contiguous amino acid residues may be both N-terminal and C-terminal to the GTP-binding domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domains are characteristic of a signature specific to GTP-binding proteins. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:

YYRGAVGALLVFDLTKHQTYAVVERWLKELYDHXEATIVVMLVGNK (SEQ ID NO: 128),

LCLQGYYRGAVGALLVFDLTKHQTYAVVERWLKELYDHXEATIVVMLVGNKMTXA RPGKCP (SEQ ID NO:88), and/or

MRXKMGNGTEEDYNFVFKVVLIGESGVGKTNLLSRFTRNEFSHDSRTTIGVEFSTRTV MLGTAAVKAQIWDTAGLERYRAITSAYYRGAVGALLVFDLTKHQTYAVVERWLKEL

YDHAEATIVVMLVGNKSDLSQAREVPTEEARMFAENNGLLFLETSALDSTNVELAFET VLKEIFAKVSKQRQNSIRTNAITSGSAQAGQEPGPGEKRACCISL (SEQ ID NO:89). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer tissue. Additionally, this gene is expressed in breast, placenta, and ovarian tumor.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including, but not limited to, prostate cancer, breast cancer, ovarian tumor. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, breast, ovary, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, or all five of the immunogenic epitopes shown in SEQ ID NO: 64 as residues: Gly-2 to Tyr-9, Phe-32 to Phe-37, Arg-133 to Glu-139, Ser-178 to Ile-185, Ala-197 to Arg-207. Polynucleotides encoding said polypeptides are also encompassed by the invention.

The HRAB-2 protein is a known GTP-binding protein. This invention relates to the novel finding of expression of HRAB-2 in prostate tissues. The tissue distribution in prostate

and prostate cancer tissue, indicates that polynucleotide and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies to the polynucleotides or polypeptides corresponding to this gene) would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as breast cancer, and ovarian cancer. The expression of HRAB-2 in the prostate tissue may indicate that polynucleotide and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies to the polynucleotides or polypeptides corresponding to this gene) of HRAB-2 could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Likewise, the expression in the breast tissue may indicate that polynucleotide and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies to the polynucleotides or polypeptides corresponding to this gene) of HRAB-2 have uses in breast neoplasia and breast cancers, such as fibroadenoma, pipillary carcinoma, ductal carcinoma, Pagetís disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma, as well as juvenile hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases. Furthermore, that

polynucleotide and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies to the polynucleotides or polypeptides corresponding to this gene) of HRAB-2 may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Polynucleotide and polypeptide fragments, agonists and/or antagonists of HRAB-2, as well as, antibodies directed against the same may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1115 of SEQ ID NO:15, b is an integer of 15 to 1129, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The translation product of this gene shares sequence homology with human neuropeptide Y (NPY) (See, e.g. Genbank Accession No. gb|K01911|HUMNPY; all references available through this accession are hereby incorporated by reference herein) one of the most abundant peptides in the mammalian nervous system. NPY is implicated in the control of feeding and the secretion of the gonadotrophin-releasing hormone.

The present invention relates to the novel finding of NPY expression in prostate and prostate cancer tissue. Preferred embodiments of the invention are polynucleotide and polypeptide fragments, agonists and/or antagonists of NPY (including antibodies directed to NPY or to fragments thereof) which can be used to treat and/or diagnose disorders of the prostate, including prostate cancer.

Included in this invention as preferred domains are pancreatic hormone family signature domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). Pancreatic hormone (PP) is a peptide synthesized in pancreatic islets of Langherhans, which acts as a regulator of pancreatic and gastrointestinal functions. A number of other active peptides are homologous to pancreatic hormone, including Neuropeptide Y (NPY). All these peptides are 36 to 39 amino acids long. Like most active peptides, their C-terminal is amidated and they are synthesized as larger protein precursors. The signature for these peptides is based on the last 17 C-terminal residues, where three positions are completely conserved. The consensus pattern is as follows: [FY]-x(3)-[LIVM]-x(2)-Y-x(3)-[LIVMFY]-x-R-x-R-[YF]. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence: YYSALRHYINLITRQRY (SEQ ID NO: 129). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. Further preferred are polypeptides comprising the pancreatic hormone family signature domain listed above, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of the amino acid sequence reference in Table I for this gene. The additional contiguous amino acid residues may be N-terminal or C- terminal to the pancreatic hormone family signature domain. Alternatively, the additional contiguous amino acid residues may be both N-terminal and C-terminal to the pancreatic hormone family signature domain, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred

polypeptide domain is a signature specific to Pancreatic hormone family proteins and homologs like NPY.

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: MLGNKRLGLSGLTSPCPCSCAWVRWPRRTPPSRTTRARTHQRRTWPDTTQRCDTTSTS SPGRDMENDLAQRH (SEQ ID NO:90), and/or MLGNKRLGLSGLTLALSLLVCLGALAEAYPSXPDNPGEDAPXEGHGQILLXAATLHQP HHQAEIWKTIXPRDTDFRPLDERKHRKCSQNSA (SEQ ID NO:91). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate or prostate cancer tissue as determined by expression analysis described in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the male reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, prostate, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two,

or all three of the immunogenic epitopes shown in SEQ ID NO: 65 as residues: Tyr-29 to Ala-42, Thr-60 to Pro-70, Glu-81 to Glu-92. Polynucleotides encoding said polypeptides are also encompassed by the invention.

The tissue distribution in prostate cancer tissue, indicates that polynucleotide and polypeptide fragments corresponding to this gene, agonists and/or antagonists of NPY (including antibodies directed to NPY or to fragments thereof) would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate polynucleotide and polypeptide fragments, agonists and/or antagonists of NPY (including antibodies directed to NPY or to fragments thereof) can be used to treat and diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 955 of SEQ ID NO:16, b is an integer of 15 to 969, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

Included in this invention as preferred domains are Cytochrome c family heme-binding domains, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). In proteins belonging to cytochrome c family, the heme group is covalently attached by thioether bonds to two conserved cysteine residues. The consensus sequence for this site is Cys-X-X-Cys-His and the histidine residue is one of the two axial ligands of the heme iron. This region is used a signature pattern for this class of proteins. The consensus pattern is as follows: C[^CPWHF][^CPWR]CH[^CFYW]. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: CNICHSL (SEQ ID NO: 131), ICNICHS (SEQ ID NO: 132), ICNICHSL (SEQ ID NO: 133). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides and polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. Further preferred are polypeptides comprising the above listed Cytochrome c family heme-binding sites, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of the amino acid sequence referenced in Table 1 for this gene. The additional contiguous amino acid residues may be N-terminal or C-terminal to the Cytochrome c family heme-binding sites. Alternatively, the additional contiguous amino acid residues may be both N-terminal and C-terminal to the Cytochrome c family heme-binding sites, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domains are characteristic of a signature specific to Cytochrome c family of proteins, however, other proteins carry this domain as well, e.g., IL9R, bone morphogenetic protein receptor type IA precursor, and laminin beta-3 chain precursor.

The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 16.

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:
TRPRVHLATVSASAAWDALGLPVRSHMQGSTRRMGVMTDVHRRFLQLLMTHGVLEE
WDVKRLQTHCYKVDRNATVDKLEDFINNINSVLESLYIEIKRGVTEDDGRPIYALVNL
ATTSISKMATDFAENELDLFRKALLIIDSETGFASSTNILNLVDQLKGKKMRKKEAEQV
LQKFVQNKWLIEKEGEFTLHGRAILEMEQYIRETPDAVKICNICHSLLIQGQSCETCGIR
MHLPCVAKYFQSNAEPRCPHCNDYWPHEIPKVFDPEKERESGVKSNKKSLRSRQH
(SEQ ID NO: 130) and/or
MQGSTRRMGVMTDVHRRFLQLLMTHGVLEEWDVKRLQTHCYKVHDRNATVDKLED
FINNINSVLESLYIEIKRGVTEDDGRPIYALVNLATTSISKMATDFAENELDLFRKALELII
DSETGFASSTNILNLVDQLKGKKMRKKEAXQVLQKFVQNKWLIEKEGEFTLHGRAILE
MEQYIRETYPDAVKICNICHSLLIQGQSCETCGIRMHLPCVAKYFQSNAEPRCPHCNDY
WPHEIPKVFDPEKERESGVLKSNKKSCGPGSISHRALLRGWLP (SEQ ID NO: 134).
Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer tissues, as determined by expression analysis described in Example 3. Additionally, this gene is expressed in breast cancer, and endometrial tumor.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including but not limited to prostate cancer, breast cancer and endometrial

tumors. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, endometrial, breast, cancerous and wounded tissues) or bodily fluids (e.g., semen, vaginal pool, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, six, seven, or all eight of the immunogenic epitopes shown in SEQ ID NO: 66 as residues: Met-1 to Arg-7, Thr-38 to Thr-49, Arg-73 to Arg-81, Leu-136 to Ala-146, Arg-181 to Asp-186, Phe-220 to Tyr-234, Phe-243 to Glu-250, Lys-255 to His-266. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate, breast cancer and endometrial tumors and, in particular, the elevated expression levels in prostate and prostate cancer indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment and diagnosis of tumors, especially breast cancer and prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate that polynucleotides and polypeptides corresponding to this gene could be used in the disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Likewise, the expression in the breast tissue may indicate that polynucleotides and polypeptides corresponding to this gene would be useful in treating, preventing, detecting and/or diagnosing breast neoplasia and breast cancers, such as fibroadenoma, pipillary carcinoma, ductal carcinoma, Pagetís disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma, as well as juvenile hypertrophy and

gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1084 of SEQ ID NO:17, b is an integer of 15 to 1098, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:

INKQMNYLFFFLTTSGLYCLSGSHGSNVKYIVLTYFNCSWSLTSPGFRDVLKGSQLWQ VTDSWEMERTKEYSSCLTFLPTADIVQARVMEELNLLASQAAPIPTSQCTAPPHLFSPLS LTSPFIMSHKSGTVGSHYNLLCHRDSIFLISNHVS (SEQ ID NO: 135);

FNCSWSLTSPGXRDVLKGSQLWQVTDSWEMERTKEYSSCLTFLPTADIVQARVMEEL NLLASQAAPIPTSQCTAPPHLFSPLSLTSPFIMSHKSGTVGSHYNLLCHRDSIFLISNHVS (SEQ ID NO: 136) and/or

FLHTFNCSWSLTSPGXRDVLKGSQLWQVTDSWEMERTKEYSSCLTFLPTADIVQARV MEELNLLASQAAPIPTSQCTAPPHLFSPLSLTSPFIMSHKSGTVGSHYNLLCHRDSIFLIS

NHVS (SEQ ID NO:93). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome X. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome X.

This gene is expressed primarily in prostate and prostate cancer tissues as determined by expression analysis described in Example 3. Additionally, this gene is expressed in colon carcinoma.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, or cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, prostate, colon, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one or both of the immunogenic epitopes shown in SEQ ID NO: 67 as residues: Glu-62 to Ser-69, Pro-99 to Thr-104. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and prostate cancer indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment, prevention,

detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated (e.g., colon cancer). The expression in the prostate tissue indicates the gene or its products could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. In addition, the tissue distribution in colon carcinoma indicates that polynucleotides and polypeptides corresponding to this gene would be useful for diagnosis, treatment and/or detection of tumors, especially of the intestine, such as, carcinoid tumors, lymphomas, cancer of the colon and cancer of the rectum, as well as cancers in other tissues where expression has been indicated. Expression in the colon tissue indicates the gene or its products would be useful for the diagnosis, treatment and/or prevention of disorders of the colon, including inflammatory disorders such as, diverticular colon disease (DCD), inflammatory colonic disease, Crohn's disease (CD), non-inflammatory bowel disease (non-IBD) colonic inflammation; ulcerative disorders such as, ulcerative colitis (UC), amebic colitis, eosinophilic colitis; noncancerous tumors, such as, polyps in the colon, adenomas, leiomyomas, lipomas, and angiomas. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1978 of SEQ ID

NO:18, b is an integer of 15 to 1992, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:

RTNLKEASDIKLEPNTLNGYKSSVTEPCPDSGEQLQPAPVLQEEELAHETAQKGEAKC HKSDTGMSKKKRQGKLVKQFAKIEESTPVHDSPGKDDAVPDLMGPHSDQGEHSGTVG VPVSYTDCAPSPVGCSVVTSDSFTKDSFRTAKSKKKRRITRYDAQLILENNSGIPKLTLR RRHDSSSKTNDQENDGMNSSKISIKLSKDHDNNNLYVAKLNNGFNSGSGSSSTKLKIQ LKRDEENRGSYTEGLHENGVCCSDPLSLLESRMEVDDYSQYEESTDDSSSSEGDEEED DYDDDFEDDFIPLPPAKRLRLIVGKDSIDIDISSRRREDQSLRLNA (SEQ ID NO: 137).

Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer as determined by expression analysis described in Example 3. Additionally, this gene is expressed in placenta, pregnant uterus and ovary tumor.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at

significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., vaginal pool, semen, amniotic fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, six, seven, eight, nine ten, or all eleven of the immunogenic epitopes shown in SEQ ID NO: 68 as residues: Met-1 to Gly-9, Thr-22 to Ala-33, Pro-40 to Gly-49, Asp-73 to Asp-96, Arg-113 to Asn-131, Leu-139 to Asn-148, Asn-155 to Thr-167, Lys-174 to Tyr-183, Glu-185 to Asn-190, Met-205 to Asp-239, Ser-264 to Ser-272. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and prostate cancer indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment and diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products can be used to treat and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from

the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1611 of SEQ ID NO:19, b is an integer of 15 to 1625, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence: EAAVSKPAGNWDVAGDERTDPSVLPA (SEQ ID NO: 138). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. The gene encoding the disclosed cDNA is believed to reside on chromosome 3. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in prostate and prostate cancer as determined by expression analysis described in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, prostate, cancerous and wounded tissues) or bodily fluids (e.g., semen,

lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in prostate and prostate carcinoma indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment and diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products can be used to treat and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Moreover, the predicted surface expression and specific prostate expression indicates that this gene would be a good target for antagonists, particularly small molecules or antibodies, which bind the polypeptides of the invention. Accordingly, preferred are antibodies and or small molecules which specifically bind an extracellular portion of the translation product of this gene. The extracellular regions can be ascertained from the information regarding the transmembrane domains as set out above. Also provided is a kit for detecting prostate cancer. Such a kit comprises in one embodiment an antibody specific for the translation product of this gene bound to a solid support. Also provided is a method of detecting prostate cancer in an individual which comprises a step of contacting an antibody specific for the translation product of this gene to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are more particularly described elsewhere herein. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional

supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2916 of SEQ ID NO:20, b is an integer of 15 to 2930, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

The translation product of this gene shares sequence homology with rabbit, murine and human alpha 1,2-exo-mannosidases which are known to be involved in N-linked oligosaccharide processing (See, e.g., Genbank Accession Nos: gi|2154997|emb|CAA73105.1|, pir|B54408|B54408, and gi|416180|emb|CAA52831.1|; all references accessible through these accessions are hereby incorporated by reference herein, for example, Bause, et al., Eur. J. Biochem. 217 (2), 535-540 (1993)). Structurally, this family of proteins are type II transmembrane proteins consisting of a short cytoplasmic polypeptide tail, a single transmembrane domain acting as a non-cleavable signal sequence and a large COOH-terminal catalytic domain containing two potential N-glycosylation sites. Two predicted N-glycosylation sites were identified within the amino acid sequence referenced in Table I for this gene using the ProSite analysis tool (Swiss Institute of Bioinformatics) which comprise the following amino acid residues: NVSG (SEQ ID NO: 140) and/or NHSD (SEQ ID NO: 141). Thus, a preferred embodiment of the invention is the catalytic domain of a novel 1,2-exo-mannosidase comprising the following amino acid sequence: AFAKSYLGDTIEGTPAGTGPEFPGRPTRPMQFAWQSYKRYAMGKNELRPLTKDGYEG

NMFGGLSGATVIDSLDTLYLMELKEEFQEAKAWVGESFHLNVSGEASLFEVNIRYIGG
LLSAFYLTGEEVFRIKAIRLGEKLLPAFNTPTGIPKGVVSFKSGNWGWATAGSSSILAEF
GSLHLEFLHLTELSGNQVFAEKVRNIRKVLRKIEKPFGLYPNFLSPVSGNWVQHHVSV
GGLGDSFYEYLIKSWLMSGKTDMEAKNMYYEALEAIETYLLNVSPGGLTYIAEWRGG
ILDHKMGHLACFSGGMIALGAEDAKEEKRAHYRELAAQITKTCHESYARSDTKLGPEA
FWFNSGREAVATQLSESYYILRPEVVESYMYLWRQTHNPIYREWGWEVVLALEKYCR
TEAGFSGIQDVYSSTPNHDNKQQSFFLAETLKYLYLLFSEDDLLSLEDWVFNTEAHPLP
VNHSDSSGRAWGRH (SEQ ID NO: 139). Moreover, fragments and variants of this polypeptide (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to this polypeptide, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding this polypeptide) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. It has been demonstrated that N-terminal deletion mutants of Man9-mannosidase are still catalytically active (See, e.g. Bieberich, et al. Eur. J. Biochem. 246 (3), 681-689 (1997)), therefore, based on the sequence similarity to the catalytic domain, the translation product of this clone is expected to share at least some biological activities with 1,2-exo-mannosidase proteins. Such activities are known in the art, some of which are described elsewhere herein. 1,2-exo-mannosidase activity may be easily assayed by techniques known in the art, such as in Bieberich et al., supra. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:

FGTSYIGGLLSAFYLTGEEVFRIKAIRLGEKLLPAFNTPTGIPKGVVSFKSGNWGWATA
GSSSILAEFGSLHLEFLHLTELSGNQVFAEKVRNIRKVLRKIEKPFGLYPNFLSPVSGNW
VQHHVSVGGLGDSFYEYLIKSWLMSGKTDMEAKNMYYEALEAXRDLLAECXSRGAD
LHCRVARGDSGPQDGAPGLFLRGHDRPWPEDAKEEKRAHYRELAAQITKTCHESYAR
SDTKLGPEASGLTPAERPWPPS (SEQ ID NO:96) and/or

MQFAWQSYKRYAMGKNELRPLTKDGYEGNMFGGLSGATVIDSLDTLYLMELKEEFQ
EAKAWVGESFHLNVSGEASLFEVNIRYIGGLLSAFYLTGEEVFRIKAIRLGEKLLPAFNT

PTGIPKGVVSFKSGNWGWATAGSSSILAEFGSLHLEFLHLTELSGNQVFAEKVRNIRKV LRKXEKPFGLY (SEQ ID NO:95). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in prostate and prostate cancer tissue as determined by expression analysis as described in Example 3. Additionally, this gene is expressed in placenta, ovary, ovarian tumor, and pineal gland.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including but not limited prostate and ovarian cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, placental, cancerous and wounded tissues) or bodily fluids (e.g., semen, vaginal pool, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, six, seven, eight, or all nine of the immunogenic epitopes shown in SEQ ID NO: 70 as residues: Ser-7 to Ala-12; Gly-14 to

Met-30, Lys-52 to Ala-58, Gly-223 to Ala-229, Glu-283 to Arg-290, Thr-303 to Gly-316, Leu-351 to Asn-357, Tyr-388 to Gln-399, His-438 to Trp-446. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and other reproductive system tissues and homology to alpha 1,2-exo-mannosidase catalytic domains indicates that polynucleotides and polypeptides corresponding to this gene would be useful for diagnosis, detection, treatment and/or prevention of disorders of the reproductive system, specifically those involving regulation of oligosaccharide processing.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2641 of SEQ ID NO:21, b is an integer of 15 to 2655, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

The gene encoding the disclosed cDNA is believed to reside on chromosome 1. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 1.

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:
MTFQFNFTIEDHLEELTPIRDGALTLDSSKELSVSESQKGEERDRKCSAEQFDLPQDHL WEHKSMENAAPSQDTDSPLSAASSRNLGATWENSPP (SEQ ID NO:97). Moreover, fragments and variants of this polypeptide (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these

polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer tissues as determined in expression analysis described in Example 3. Additionally, this gene is expressed in fetal liver/spleen and germinal center B cells.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the hematopoietic system or male reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, hematopoietic, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hematopoietic, prostate, developing, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, six, seven, eight, or all nine of the immunogenic epitopes shown in SEQ ID NO: 71 as residues: Glu-37 to Glu-51, Phe-53 to His-59, Ala-70 to Ser-77, Ser-83 to Ser-96, Pro-106 to Lys-111, Thr-236 to Asn-245, Lys-250 to Val-257, Pro-299 to Asn-304, Glu-342 to Phe-347. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and proliferating tissues indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment and diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products can

be used to treat, prevent, detect, and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Alternatively, the tissue distribution in fetal liver/spleen and germinal center B cells indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the diagnosis and/or treatment of hematopoietic disorders. Elevated levels of expression of this gene product in germinal center B cells suggests that it may play a role in the survival, proliferation, and/or growth of B cells. The gene product may be be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. More generally, as evidenced by expression in fetal liver/spleen, this gene may play a role in the survival, proliferation, and/or differentiation of hematopoietic cells in general, and may be of use in augmentation of the numbers of stem cells and committed progenitors. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1559 of SEQ ID NO:22, b is an integer of 15 to 1573, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

The translation product of this gene shares sequence homology with Mad2B protein which is thought to be a necessary component of mitotic checkpoints (See, e.g., Genbank Accession No. gi|4835900|gb|AAD30290.1|AF139365_1; all references available through this accession are hereby incorporated by reference herein). Aneuploidy is a characteristic of the majority of human cancers, and recent work has suggested that mitotic checkpoint defects play a role in its development. Based on the sequence similarity, the translation product of this gene is expected to share at least some biological activities with Mad2B proteins. Such activities are known in the art, some of which are described elsewhere herein. Such activities may be assayed according to well known techniques such as that described in Wang, et al. I Chuan Hsueh Pao. 26(2):112-8(1999).

Preferred novel polynucleotides of the invention comprise, or alternatively consist of, the nucleotide sequence from about 561 to about 1645 of the primary nucleotide sequence referenced in Table I for this gene. Moreover, fragments and variants of these polynucleotides (such as, for example, fragments as described herein, polynucleotides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to this nucleotide sequence, or polynucleotides which hybridize, under stringent conditions, to this nucleotide sequence) are encompassed by the invention. Polypeptides encoded by these polynucleotides are also preferrered. Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancers as determined by expression analysis described herein in Example 3. This gene is also expressed in testes, colon carcinoma and breast cancer.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for

diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system including, but not limited to, prostate cancer and breast cancer; and cancers. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, vaginal pool, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, or all four of the immunogenic epitopes shown in SEQ ID NO: 72 as residues: Gln-43 to Asn-49, Glu-59 to Gln-65, Lys-90 to Val-95, Glu-205 to Ser-211. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and prostate cancer and homology to human Mad2B protein indicates that polynucleotides and polypeptides corresponding to this gene, fragments thereof, agonists and antagonists (including antibodies directed to the protein encoded by this gene or to fragments thereof) corresponding to this gene would be useful for the treatment, prevention, detection, and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated (e.g., breast, colon and testes). Mad2B protein homologs are thought to be necessary components of the mitotic checkpoint. Defects of this gene lead to aneuploidy, a characteristic of the majority of human cancers. The expression in the prostate tissue may indicate the gene or its products, fragments, agonist, and/or antagonists thereof (including antibodies directed to the protein encoded by this gene or fragments thereof) could be used to treat and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with

systemic or reproductive functions. Likewise, the expression in the breast tissue may indicate that polynucleotides and/or polypeptides corresponding to this gene, fragments thereof, and agonists and/or antagonists thereof (including antibodies to the protein encoded by this gene, or fragments thereof) would be useful in the treatment, prevention, detection and/or diagnosis of breast neoplasia and breast cancers, such as fibroadenoma, pipillary carcinoma, ductal carcinoma, Pagetís disease, medullary carcinoma, mucinous carcinoma, tubular carcinoma, secretory carcinoma and apocrine carcinoma, as well as juvenile hypertrophy and gynecomastia, mastitis and abscess, duct ectasia, fat necrosis and fibrocystic diseases. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1984 of SEQ ID NO:23, b is an integer of 15 to 1998, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

This clone encodes a portion of a pseudoautosomal gene called PGPL (See, e.g., Genbank Accession No. gi|2765411|emb|CAA74749.1|; all references available through this accession are hereby incorporated by reference herein), which escapes X inactivation and has a functional homologue on the Y chromosome.

The present invention relates to the novel finding of PGPL expression in prostate and prostate cancer tissue. Preferred embodiments of the invention are polynucleotide and polypeptide fragments, variants or derivatives corresponding to this gene (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides); and agonists and/or antagonists (such as, for example, antibodies or fragments thereof which bind polynucleotides or polypeptides corresponding to this gene) which can be used to treat, prevent, detect and/or diagnose disorders of the prostate, including prostate cancer. Included in this invention as preferred domains are ATP/GTP-binding site motifs, which were identified using the ProSite analysis tool (Swiss Institute of Bioinformatics). From sequence comparisons and crystallographic data analysis it has been shown that an appreciable proportion of proteins that bind ATP or GTP share a number of more or less conserved sequence motifs. The best conserved of these motifs is a glycine-rich region, which typically forms a flexible loop between a beta-strand and an alpha-helix. This loop interacts with one of the phosphate groups of the nucleotide. This sequence motif is generally referred to as the 'A' consensus sequence or the 'P-loop'. The consensus pattern is as follows: [AG]-x(4)-G-K-[ST]. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: GYTNCGKT (SEQ ID NO: 142), VGYTNCGKTT (SEQ ID NO: 143), and/or VVGYTNCGKTTL (SEQ ID NO: 144). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. Further preferred are polypeptides comprising the ATP/GTP-binding site motifs listed above, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of the amino acid sequence reference in Table I for this gene. The additional contiguous amino acid residues may be N-

terminal or C- terminal to the ATP/GTP-binding site motifs. Alternatively, the additional contiguous amino acid residues may be both N-terminal and C-terminal to the ATP/GTP-binding site motifs, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. The above preferred polypeptide domain is a signature specific to ATP/GTP-binding related proteins.

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:

PHRVDTRRRDPVPRSRALSHGTGRVGAAAGESSRAPRCWSGSRPRAPADPPRHRPLLC LSRRGSPPHHLGCLLGESFMQLQQRLLREKEAKIRKALDRLRKKRHLLRRQRTRREFP VISVVGYTNCGKTTLIKALTGDAAIQPRDQLFATLDVTAHAGTLPSRMTVLYVDTIGFL SQLPHGLIESFSATLEDVAHSDLILHVRDVSHPEAELQKCSVLSTLRGLQLPAPLLDSMV EVHNKVDLVPGYSPTEPNVVPVSALRGHGLQELKLSSMRRFXRRRGDRSSLSV (SEQ ID NO:98). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate or prostate cancer tissue as determined by expression analysis described in Example 3.

Polynucleotide and polypeptide fragments, variants or derivatives corresponding to this gene (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides); and agonists and/or antagonists (such as, for example, antibodies or fragments thereof which bind polynucleotides or polypeptides corresponding to this gene) would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and

conditions which include but are not limited to: disorders of the male reproductive system, including but not limited to prostate cancer. Similarly, polypeptide fragments and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in prostate cancer tissue, indicates that polynucleotide and polypeptide fragments, variants or derivatives corresponding to this gene (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides); and agonists and/or antagonists (such as, for example, antibodies or fragments thereof which bind polynucleotides or polypeptides corresponding to this gene) would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate that polynucleotide and polypeptide fragments, variants or derivatives corresponding to this gene (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides); and agonists and/or antagonists (such as, for example, antibodies or fragments thereof which bind polynucleotides or polypeptides corresponding to this gene) can be used to treat and diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or

as hormones or factors with systemic or reproductive functions. Additionally the finding of PGPL in prostate cancer tissues indicates that this gene would be a good target for antagonists, particularly small molecules or antibodies, which block binding of the receptor by its cognate ligand(s). Accordingly, preferred are antibodies and or small molecules which specifically bind an extracellular portion of the translation product of this gene. The extracellular regions can be ascertained by methods known in the art. Also provided is a kit for detecting testicular cancer. Such a kit comprises in one embodiment an antibody specific for PGPL bound to a solid support. Also provided is a method of detecting prostate cancer in an individual which comprises a step of contacting an antibody specific for PGPL or a fragment thereof to a bodily fluid from the individual, preferably serum, and ascertaining whether antibody binds to an antigen found in the bodily fluid. Preferably the antibody is bound to a solid support and the bodily fluid is serum. The above embodiments, as well as other treatments and diagnostic tests (kits and methods), are more particularly described elsewhere herein. Furthermore, polynucleotide and polypeptide fragments, variants or derivatives corresponding to this gene (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides); and agonists and/or antagonists (such as, for example, antibodies or fragments thereof which bind polynucleotides or polypeptides corresponding to this gene) may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from

the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 978 of SEQ ID NO:24, b is an integer of 15 to 992, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:

RHHDRSPLSDPLLPETLLAPPDPPGLWPAAPLSLRRRGSAVTHQRASGRGWGGGAGMS LPLRAPAPRLERRPAGPPADVFLVPKRVVRASRPLRDLRASHRAPRTQRAWSSPLTPSP AGTHAGSTHSAPPPNFWERTPGSAQPLAFQKPLYAYLIFVIGDEPSLLSPFPHTHQSPLAI PSPSASPPPSCAPAPHSPPPIGLALACKSRRWPRAQPSRMSPGPPLWERRQSYWPLTRPL GPRARQAFESTCSSPESRPRPCLPHRRPQSTLPQL (SEQ ID NO: 145); ALWAGAGGFEGLSSTRAQRSCQWPVALPPFPERGSRGHPGRLGPGPPSALAS (SEQ ID NO: 146); and/or

PPPIGLALACKSRRWPRAQPSRMSPGPPLWERRQSYWPLTRPLGPRARQAFESTCSSPE SXPXRASHTAADLRAPCLNCEFFLGNPLKRKGYQS (SEQ ID NO:99). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer tissue as determined by expression analysis described in Example 3. Additionally, this gene is expressed in prostate, healing groin, and thyroid tumor.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for

diagnosis of diseases and conditions which include but are not limited to: disorders of the endocrine and reproductive systems, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and/or endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., endocrine, reproductive, prostate, cancerous and wounded tissues) or bodily fluids (e.g., lymph, semen, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, six, seven, eight, or all nine of the immunogenic epitopes shown in SEQ ID NO: 74 as residues: Pro-11 to Gly-19, Arg-36 to Arg-42, Ala-47 to Trp-54, Ala-67 to Ser-86, Pro-139 to Pro-144, Lys-153 to Pro-167, Glu-173 to Tyr-178, Thr-196 to Cys-207, Pro-209 to Thr-217. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and thyroid tumor indicate that polynucleotides and polypeptides corresponding to this gene would be useful for the detection, diagnosis, treatment, and/or prevention of various endocrine disorders and cancers. Briefly, polynucleotides and polypeptides corresponding to this gene could be used for the detection, treatment, and/or prevention of Addison's disease, Cushing's Syndrome, and disorders and/or cancers of the pancreas (e.g., diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g., hyper-, hypothyroidism), parathyroid (e.g., hyper-, hypoparathyroidism), hypothallamus, and testes. Similarly, the tissue distribution in prostate cancer tissue, indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic

prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1912 of SEQ ID NO:25, b is an integer of 15 to 1926, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: KGIMLCWFCFFVLLFFFFFLFCFLVMWLKCNSFFFGTYFCQLKTRRAQLFF (SEQ ID NO: 147); and/or MLCWFCFFVLLFFFFFLFCFLVMWLKCNSFFLGHISAN (SEQ ID NO:100). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the

invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer as determined by expression anlaysis described in Example 3. Additionally, this gene is expressed in a variety of tumor or proliferating tissues, including, ovary tumor, colon carcinoma, and placenta.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including developmental disorders, prostate cancer, and ovarian tumor. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, ovary, breast, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., vaginal pool, lymph, semen, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in prostate, ovary tumor, colon cancer and placenta indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. The tissue distribution in colon and colon cancer indicates that polynucleotides and polypeptides corresponding to this gene would be useful for diagnosis, treatment, prevention and/or detection of tumors, especially of the intestine, such as, carcinoid

tumors, lymphomas, cancer of the colon and cancer of the rectum, as well as cancers in other tissues where expression has been indicated. In addition, expression in the colon tissue indicates the gene or its products would be useful for the diagnosis, treatment and/or prevention of disorders of the colon, including inflammatory disorders such as, diverticular colon disease (DCD), inflammatory colonic disease, Crohn's disease (CD), non-inflammatory bowel disease (non-IBD) colonic inflammation; ulcerative disorders such as, ulcerative colitis (UC), amebic colitis, eosinophilic colitis; noncancerous tumors, such as, polyps in the colon, adenomas, leiomyomas, lipomas, and angiomas. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2528 of SEQ ID NO:26, b is an integer of 15 to 2542, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:

MKTGGKHSVIRYFSNIKTTKTNDKNVYFYTPAYRVSFRVYEYLNLLISVLMKAELNRE S (SEQ ID NO: 148). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%,

97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer as determined by expression analysis described in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the male reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, prostate, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, or all three of the immunogenic epitopes shown in SEQ ID NO: 76 as residues: Met-1 to Lys-6, Ile-16 to Asn-25, Asn-57 to Ser-62. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and prostate cancer indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products can be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous

prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1625 of SEQ ID NO:27, b is an integer of 15 to 1639, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14..

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:

PGKPKSAHFPPCCMFSVLCLCVCARQRDRLFVKSASCLGIFVSHLAVSSRTIQLAFQAW R (SEQ ID NO: 149) and/or

WMSEYXQWVFLISLRICLRVHYQGSGTRXHSLHQFLRVL (SEQ ID NO: 150). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by

the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer as determined by expression analysis described in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the male reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, prostate, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in prostate and prostate cancer indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products can be used to treat and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1802 of SEQ ID NO:28, b is an integer of 15 to 1816, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene encodes a fragment of the human monoamine oxidase A protein. MOA is thought to play a vital role in the metabolism of biogenic amines in the central nervous system and in peripheral tissues (See, e.g., Bach, et al., Proc. Natl. Acad. Sci. U.S.A. 85:4934-4938 (1988)).

The present invention relates to the novel finding of MOA expression in prostate and prostate cancer tissue. Preferred embodiments of the invention are polynucleotide and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies which bind to the polynucleotides or polypeptides corresponding to this gene) of MOA which can be used to treat, prevent, detect and/or diagnose disorders of the prostate, including prostate cancer. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: CELYAKVLGSQEALHPVHYEEKNWC (SEQ ID NO: 153), CELYAKVLGSQEALHPHYEEKNWCEEQYSGGCY (SEQ ID NO: 152), and/or RKKKICELYAKVLGSQEALHPHYEEKNWCEEQYSGGC (SEQ ID NO: 151). Moreover,

fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention. Further preferred are polypeptides comprising the fragments listed above, and at least 5, 10, 15, 20, 25, 30, 50, or 75 additional contiguous amino acid residues of MOA. The additional contiguous amino acid residues may be N-terminal or C-terminal to the fragments. Alternatively, the additional contiguous amino acid residues may be both N-terminal and C-terminal to the fragments, wherein the total N- and C-terminal contiguous amino acid residues equal the specified number. In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:

MGAVIKCMMYYKEAFWKKKDYCGCMIIEDEDAPISITLDDTKPDGSLPAIMGFILARK ADRLAKLHKEIRKKKICELYAKVLGSQEALHPVHYEEKNWCEEQYSGGCYTAYFPPGI MTQYGRVIRQPVGRIFFAGTETATKWSGYMEGAVEAGERAAREVLNGLGKVTEKDIW VQEPESKDVPAVEITHTFWERNLPSVSGLLKIIGFSTSVTALGFVLYKYKLLPRS (SEQ ID NO:103). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer tissues. Additionally, this gene is expressed in fetal liver/spleen and placenta.

Polynucleotides and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a

polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies which bind to the polynucleotides or polypeptides corresponding to this gene) of MOA would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including, but not limited to, prostate cancer. Similarly, polypeptide fragments and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, placenta, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, six, seven, or all eight of the immunogenic epitopes shown in SEQ ID NO: 78 as residues: Pro-34 to Arg-39, Asn-80 to Phe-87, Lys-184 to Cys-189, Asp-206 to Gly-212, His-233 to Lys-240, Tyr-260 to Gly-273, Thr-305 to Gly-311, Glu-343 to Asp-348. Polynucleotides encoding said polypeptides are encompassed by the invention.

The novel finding of MOA expression in prostate and prostate cancer tissue indicates that polynucleotides and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies which bind to the polynucleotides or polypeptides corresponding to this gene) of MOA would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer. The expression in the prostate tissue may indicate that polynucleotides and polypeptide fragments,

variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies which bind to the polynucleotides or polypeptides corresponding to this gene) of MOA can be used to treat and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, polynucleotides and polypeptide fragments, variants or derivatives (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), agonists and/or antagonists (such as, for example, antibodies which bind to the polynucleotides or polypeptides corresponding to this gene) of MOA may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to their use as a nutritional supplement. Polynucleotide and polypeptide fragments, agonists and/or antagonists of MOA, as well as, antibodies directed against the same may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3133 of SEQ ID NO:29, b is an integer of 15 to 3147, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

The gene encoding the disclosed cDNA is believed to reside on chromosome 16. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 16.

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, the following amino acid sequence:

FFVIPSSGSICFCSLVTVLMFNCCTLKPKSVTMHTVTKVLGLQSCLLYKENFKCCCKLT SYTILNFLSSPLFLPTNGIIMLA (SEQ ID NO:104). Moreover, fragments and variants of this polypeptide (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to this polypeptide, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding this polypeptide) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed in prostate and prostate cancer tissue as determined by expression analysis described in Example 3. Additionally, this gene is expressed in fetal liver and spleen, glioblastoma and tongue.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the prostate, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, vaginal pool, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken

from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of the immunogenic epitopes shown in SEQ ID NO: 79 as residues: Leu-38 to Gly-44. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate indicates that the polynucleotides and polypeptides corresponding to this gene would be useful for detection, diagnosis, treatment and/or prevention of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated (e.g., glioblastoma). The expression in the prostate tissue may indicate that the polynucleotides and polypeptides corresponding to this gene can be used in the disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1095 of SEQ ID NO:30, b is an integer of 15 to 1109, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

The gene encoding the disclosed cDNA is believed to reside on chromosome 5. Accordingly, polynucleotides related to this invention would be useful as a marker in linkage analysis for chromosome 5.

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: QLLLLPPKAPRNPFLPCPGSRTPGYIWKVEMWGSCVLEYYVSPPSAVFSEHVCCPWWE RGHCAVVHRCLSFTVGLSVCLSFLSAAQMENNYLLHWRERKSLRIPKGTLA (SEQ ID NO: 154) and/or MWGSCXLEYYVSPPSAVFSEHVCCPWWERGHCAVVHRCLSFTVGLSVCLSFLSAAQM ENNYLLHWRERKSLRIPKGTLA (SEQ ID NO:105). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed in prostate and prostate cancer tissue as determined by expression analysis described in Example 3. Additionally, the gene is expressed in fetal heart, lung, and liver and spleen.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types

(e.g., prostate, developing tissues, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, vaginal pool, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise the immunogenic epitopes shown in SEQ ID NO: 80 as residues: Arg-66 to Leu-71. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and prostate cancer tissues indicates that polynucleotides and polypeptides corresponding to this gene would be useful in diagnosis, treatment and/or prevention of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate that polynucleotides and polypeptides corresponding to this gene could be used to treat, prevent, detect and/or diagnose the disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Additionally, the expression within embryonic tissue and other cellular sources marked by proliferating cells suggests that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31

and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2310 of SEQ ID NO:31, b is an integer of 15 to 2324, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group: DEVSSKEGSMCPSSLHLAAGIVDITGALAAVSRGSKPHPKSKAD (SEQ ID NO: 155); and/or MCPFSSLHLAAGIVDITGALAAVSRGSKPHPKSKAD (SEQ ID NO:110). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer as determined by expression analysis described in Example 3. Additionally, this gene is expressed in fetal liver/spleen, pregnant uterus, ovary and prostate carcinoma.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for

differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive or hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, hematopoietic, prostate, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of the immunogenic epitopes shown in SEQ ID NO: 81 as residues: Arg-25 to Asp-36. Polynucleotides encoding said polypeptides are encompassed by the invention.

The tissue distribution in prostate and prostate carcinoma indicates that polynucleotides and/or polypeptides corresponding to this gene, fragments, agonists and/or antagonists thereof (e.g., including but not limited to antibodies which bind the encoded polypeptide) would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate polynucleotides and/or polypeptides corresponding to this gene, fragments, agonists and/or antagonists thereof (e.g., including but not limited to antibodies which bind the encoded polypeptide) would be useful in treating, detecting, preventing and/or diagnosing disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Additionally, the tissue distribution in fetal liver/spleen indicates the polynucleotides and/or polypeptides of the invention, fragments, agonists and/or antagonists thereof (e.g., including but not limited to antibodies which bind the encoded polypeptide) would be useful for the treatment, prevention, detection and diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Representative uses are

described in the "Immune Activity" and "Infectious Disease" sections below, in Example 11, 13, 14, 16, 18, 19, 20, and 27, and elsewhere herein. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The polynucleotides and/or polypeptides corresponding to this gene, fragments, agonists and/or antagonists thereof (e.g., including but not limited to antibodies which bind the encoded polypeptide) may also be involved in lymphopoiesis, and therefore, would be useful in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, polynucleotides and/or polypeptides corresponding to this gene, fragments, agonists and/or antagonists thereof (e.g., including but not limited to antibodies which bind the encoded polypeptide) may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1241 of SEQ ID NO:32, b is an integer of 15 to 1255, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

This gene is expressed primarily in prostate and prostate cancer as determined by expression analysis described herein in Example 3.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of disorders of the male reproductive system, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., reproductive, prostate, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in prostate and prostate cancer indicates that polynucleotides and polypeptides corresponding to this gene would be useful for the treatment, prevention, detection and/or diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate the gene or its products could be used to treat, prevent, detect and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33

and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 520 of SEQ ID NO:33, b is an integer of 15 to 534, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

In specific embodiments, polypeptides of the invention comprise, or alternatively consist of, an amino acid sequence selected from the group:
GQRQALCPQLILEASRLCEVSTSQHLCSSFEASNCLGKRDREMEAWIRANQPAFLVWR STWPFPWAQGHLKHCPVKLVLGCPCAWRVLKLTFQIPREQGEISRMSIAAKKCLGGLP LLTPHLAADQHSILNTLRAPSMAFDRTKSPGVVTENRSCAVTAMFPPGRQKLKSPKRT SFSSAADEWHRY (SEQ ID NO: 156),
VASILKAAPNRQILPLFLKHHHVGEPSEGWATSQDSLLGGLGYLGVLPHNVQGDIVTK LKRLCCFSVMSAGIKAVSAPCGASCGICSSPYPHNSGAQGPGLV (SEQ ID NO: 157), and/or IVTKLKRLCCFSVMSAGIKAVSAPCGASCGICSSPYPHNSGAQ (SEQ ID NO:108). Moreover, fragments and variants of these polypeptides (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides) are encompassed by the invention. Antibodies that bind polypeptides of the invention and polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in prostate and prostate cancer as determined by expression analysis described in Example 3. Additionally, this gene is expressed in prostate carcinoma, osteoclastoma, bone marrow and brain.

Polynucleotides and polypeptides of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive, hematopoietic and/or skeletal systems. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and/or hematopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, or both of the immunogenic epitopes shown in SEQ ID NO: 83 as residues: Pro-53 to Glu-58, Pro-121 to Ser-133. Polynucleotides encoding said polypeptides are also encompassed by the invention.

The tissue distribution in prostate and prostate carcinoma indicates that polynucleotides and polypeptides corresponding to this gene, fragments, agonists and antagonists thereof (e.g., including but not limited to antibodies directed to the polypeptides encoded by this clone) would be useful for the treatment and diagnosis of tumors, especially prostate cancer, as well as cancers of other tissues where expression has been indicated. The expression in the prostate tissue may indicate that polynucleotides and polypeptides corresponding to this gene, fragments, agonists and antagonists thereof (e.g., including but not limited to antibodies directed to the polypeptides encoded by this clone) would be useful in treating, preventing, detecting and/or diagnosing disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Additionally, the tissue distribution in bone marrow indicates that polynucleotides

and polypeptides corresponding to this gene, fragments, agonists and antagonists thereof (e.g., including but not limited to antibodies directed to the polypeptides encoded by this clone) would be useful for the treatment, prevention, detection and/or diagnosis of hematopoietic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. Briefly, the uses include bone marrow cell ex-vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The polynucleotides and polypeptides corresponding to this gene, fragments, agonists and antagonists thereof (e.g., including but not limited to antibodies directed to the polypeptides encoded by this clone) may also be involved in lymphopoiesis, and therefore, would be useful in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, polynucleotides and polypeptides corresponding to this gene, fragments, agonists and antagonists thereof (e.g., including but not limited to antibodies directed to the polypeptides encoded by this clone) may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1742 of SEQ ID NO:34, b is an integer of 15 to 1756, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

The polypeptide of the invention is Melanoma Associated Antigen-10 (MAGE-10). The family of melanoma associated antigen (MAGE) proteins are known tumor antigens. MAGE-10 was shown to be expressed in testis, blood lymphocytes, and placental tissue (See, e.g., De Plaen et al., Immunogenetics 40:360-369 (1994)).

This invention relates to the novel finding of MAGE-10 in prostate and prostate cancer. Preferred embodiments of the invention are varients (such as, for example, fragments as described herein, polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to these polypeptides, or polypeptides encoded by a polynucleotide which hybridizes, under stringent conditions, to the polynucleotide encoding these polypeptides), derivatives, fragments, agonists, and/or antagonists of the MAGE-10 protein (such as, for example, antibodies which bind MAGE-10 or fragments thereof) used to treat, prevent, detect and/or diagnose disorders of the prostate, including but not limited to, prostate cancer.

This gene is expressed primarily in prostate and prostate cancer tissue as determined by expression analysis described in Example 3. Additionally, this gene is expressed in fetal liver/spleen, and placenta.

Polypeptide fragments, derivatives, variants, agonists and/or antagonists of the invention would be useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include but are not limited to: disorders of the reproductive system, specifically prostate disorders, including but not limited to prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides would be useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., prostate, reproductive, cancerous and wounded tissues) or bodily fluids (e.g., semen, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or sample taken from an

individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred polypeptides of the present invention comprise, or alternatively consist of one, two, three, four, five, six, seven, eight, nine ten, eleven or all twelve of the immunogenic epitopes shown in SEQ ID NO: 84 as residues: Met-1 to Arg-9, Glu-13 to Gln-23, Asp-37 to Thr-44, Ser-55 to Ser-62, Thr-70 to Ser-86, Asp-103 to Ser-118, Pro-124 to Pro-130, Glu-193 to Gly-198, Glu-266 to Leu-271, Asp-274 to Arg-293, Asn-319 to Ser-325, Leu-334 to Gln-342. Polynucleotides encoding said polypeptides are encompassed by the invention.

The finding of a known tumor antigen, MAGE-10, expressed in prostate and prostate cancer tissues, indicates that polynucleotide and polypeptide varients, derivatives, fragments, agonists, and/or antagonists of MAGE-10 would be useful for the treatment and diagnosis of prostate disorders, specifically prostate cancer. The expression in the prostate tissue may indicate that polynucleotide and polypeptide varients, derivatives, fragments, agonists, and/or antagonists of MAGE-10 can be used to treat and/or diagnose disorders of the prostate, including inflammatory disorders, such as chronic prostatitis, granulomatous prostatitis and malacoplakia, prostatic hyperplasia and prostate neoplastic disorders, including adenocarcinoma, transitional cell carcinomas, ductal carcinomas, squamous cell carcinomas, or as hormones or factors with systemic or reproductive functions. Furthermore, that polynucleotide and polypeptide varients, derivatives, fragments, agonists, and/or antagonists of MAGE-10 may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to their use as a nutritional supplement. Polynucleotide and polypeptide varients, derivatives, fragments, agonists, and/or antagonists of MAGE-10, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the prostate.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention.

To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1531 of SEQ ID NO:35, b is an integer of 15 to 1545, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

Table I.

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HHEPE84	PTA-416 7/23/99	pCMVSport 3.0	11	1200	1	1200	62	62	60	1	12	13	54
1	HHEPE84	PTA-416 7/23/99	pCMVSport 3.0	36	1177	72	1177	2		85				69
2	HTXFS09	PTA-416 7/23/99	Uni-ZAP XR	12	1106	1	1106	181	181	61	1	22	23	117
3	HSPAF01	PTA-497 8/11/99	pSport1	13	887	1	887	22	22	62	1	20	21	183
3	HSPAF01	PTA-497 8/11/99	pSport1	37	1173	1	1173	1		86				95
4	HWHHB29	PTA-1620 4/3/2000	pCMVSport 3.0	14	1918	1	1601	128	128	63	1	60	61	205
4	HWHHB29	PTA-1620 4/3/2000	pCMVSport 3.0	38	1927	1	1927	138	138	87	1	60	61	181
5	HLWBY67	PTA-416 7/23/99	pCMVSport 3.0	15	1129	1	1129	256		64				213
5	HLWBY67	PTA-416 7/23/99	pCMVSport 3.0	39	532	1	532	244		88				61
5	HLWBY67	PTA-416 7/23/99	pCMVSport 3.0	40	1129	1	1129	223		89				217
6	HPWBU56	PTA-497 8/11/99	Uni-ZAP XR	16	969	1	969	328	328	65	1	28	29	97
6	HPWBU56	PTA-497 8/11/99	Uni-ZAP XR	41	553	1	553	72		90				72

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
6	HPWBU56	PTA-497 8/11/99	Uni-ZAP XR	42	599	6	599	84	84	91	1	28	29	91
7	HCFBC32	PTA-416 7/23/99	pSport1	17	1098	1	1098	81		66				266
7	HCFBC32	PTA-416 7/23/99	pSport1	43	1077	1	1077	83		92				277
8	HSLJW05	PTA-416 7/23/99	Uni-ZAP XR	18	1992	1	1992	128		67				149
8	HSLJW05	PTA-416 7/23/99	Uni-ZAP XR	44	1904	1	1904	196		93				122
9	HJAAT54	PTA-416 7/23/99	pBluescript SK-	19	1625	1	1625	195		68				277
9	HJAAT54	PTA-416 7/23/99	pBluescript SK-	45	1625	1	1625	3		94				341
10	HMEFI17	PTA-498 8/11/99	Lambda ZAP II	20	2930	1	2930	213	213	69	1	68	69	94
11	HDPQE64	PTA-416 7/23/99	pCMVSport 3.0	21	2655	1	2655	556		70				449
11	HDPQE64	PTA-416 7/23/99	pCMVSport 3.0	46	593	1	593	2		95				197
11	HDPQE64	PTA-416 7/23/99	pCMVSport 3.0	47	1792	1	1792	1790		96				254
12	HISAN67	PTA-416 7/23/99	pSport1	22	1573	1	1573	388		71				372
12	HISAN67	PTA-416 7/23/99	pSport1	48	785	1	785	376		97				97

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
13	HMWJD58	PTA-416 7/23/99	Uni-ZAP XR	23	1998	1	1998	702		72				211
14	HMQAI69	PTA-497 8/11/99	Uni-ZAP XR	24	992	1	992	1		73				219
14	HMQAI69	PTA-497 8/11/99	Uni-ZAP XR	49	1433	408	1433	64		98				288
15	HRAAA23	PTA-416 7/23/99	pCMVSport 3.0	25	1926	1	1926	835		74				221
15	HRAAA23	PTA-416 7/23/99	pCMVSport 3.0	50	652	1	652	650		99				94
16	HOGCK09	PTA-416 7/23/99	pCMVSport 2.0	26	2542	1	2542	217		75				48
16	HOGCK09	PTA-416 7/23/99	pCMVSport 2.0	51	2541	1	2541	217		100				38
17	HPRAT22	PTA-416 7/23/99	Uni-ZAP XR	27	1639	1	1639	98		76				63
17	HPRAT22	PTA-416 7/23/99	Uni-ZAP XR	52	701	1	701	96		101				60
18	HTXDT74	PTA-416 7/23/99	Uni-ZAP XR	28	1816	1	1816	100		77				36
18	HTXDT74	PTA-416 7/23/99	Uni-ZAP XR	53	375	1	375	138		102				40
19	HDPRJ46	PTA-416 7/23/99	pCMVSport 3.0	29	3147	380	3147	10		78				395
19	HDPRJ46	PTA-416 7/23/99	pCMVSport 3.0	54	1146	1	1146	149		103				228

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
20	HEQAN39	PTA-416 7/23/99	pCMVSport 3.0	30	1109	1	1109	466		79				45
20	HEQAN39	PTA-416 7/23/99	pCMVSport 3.0	55	2299	1298	2299	1044		104				82
21	HFKFN13	PTA-416 7/23/99	Uni-ZAP XR	31	2324	1	2324	698	698	80	1	55	56	79
21	HFKFN13	PTA-416 7/23/99	Uni-ZAP XR	56	2259	15	2259	595		105				79
22	HPFDD04	PTA-416 7/23/99	Uni-ZAP XR	32	1255	1	1255	619		81				36
22	HPFDD04	PTA-416 7/23/99	Uni-ZAP XR	57	1325	65	1325	673		106				36
23	HPWDE02	PTA-416 7/23/99	Uni-ZAP XR	33	534	1	534	102		82				35
23	HPWDE02	PTA-416 7/23/99	Uni-ZAP XR	58	832	161	832	209		107				35
24	HPIBS12	PTA-416 7/23/99	Uni-ZAP XR	34	1756	1	1756	1030	1030	83	1	38	39	144
24	HPIBS12	PTA-416 7/23/99	Uni-ZAP XR	59	132	1	132	130		108				43
25	HPIAD06	PTA-416 7/23/99	Uni-ZAP XR	35	1545	1	1545	345		84				369

Table 1 summarizes some of the prostate cancer antigen polynucleotides encompassed by the invention and further summarizes certain characteristics of these polynucleotides and the polypeptides encoded thereby.

Column 1 in Table 1 summarizes the information corresponding to each "Gene No." described above. The second column in Table 1, provides a unique "cDNA Clone ID" identification for each prostate cancer antigen. The prostate cancer antigen nucleotide sequence identified in column five as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA Clone ID" identified in column two of Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

Each cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in column three, "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. Column four, "Vector," refers to the type of vector contained in the cDNA Clone ID.

Column six of Table 1, "Total NT Seq.," provides the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." (column seven) and the "3' NT of Clone Seq." (column eight) of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the preferred start of translation is identified in column nine as "5' NT of Start Codon."

The preferred translated amino acid sequence, is identified in column ten as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention. The amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified in column fourteen as "Last AA of ORF." Sequence may have a signal peptide

SEQ ID NO:X (where X may be any of the polynucleotide sequences disclosed in the sequence listing) and the translated SEQ ID NO:Y (where Y may be any of the polypeptide sequences disclosed in the sequence listing) are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID

NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or a cDNA contained in ATCC Deposit No.:Z. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling immediate applications in chromosome mapping, linkage analysis, tissue identification and/or typing, and a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the prostate cancer polypeptides, or fragments thereof, and/or to the polypeptides encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1 (ATCC Deposit No.:Z). The nucleotide sequence of the cDNA contained in each ATCC Deposit No.:Z can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

Also provided in Table 1 is the name of the vector which contains the cDNA plasmid. Each vector is routinely used in the art. The following additional information is provided for convenience.

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and

5,286,636), pBluescript (pBS) (Short, J. M. et al., *Nucleic Acids Res. 16:*7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., *Nucleic Acids Res. 17:*9494 (1989)) and pBK (Alting-Mees, M. A. et al., *Strategies 5:*58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Phagemid pBS may be excised from the Lambda Zap and Uni-Zap XR vectors, and phagemid pBK may be excised from the Zap Express vector. Both phagemids may be transformed into *E. coli* strain XL-1 Blue, also available from Stratagene.

Vectors pSport1, pCMVSport 1.0, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, also available from Life Technologies. See, for instance, Gruber, C. E., et al., *Focus 15:*59 (1993). Vector lafmid BA (Bento Soares, Columbia University, New York, NY) contains an ampicillin resistance gene and can be transformed into *E. coli* strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, available from Life Technologies. See, for instance, Clark, J. M., *Nuc. Acids Res. 16:*9677-9686 (1988) and Mead, D. *et al., Bio/Technology 9:* (1991).

The present invention also relates to the prostate cancer antigen genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone (ATCC Deposit No.:Z). The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include, but are not limited to, preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are allelic variants, orthologs, and/or species homologs. Procedures known in the art can be used to obtain full-length genes, allelic variants, splice variants, full-length coding portions, orthologs, and/or species homologs of genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, and/or the deposited cDNA clone (ATCC Deposit No.:Z), using information from the sequences disclosed herein or the clones deposited with the ATCC. For example, allelic variants and/or species homologs may be isolated and identified by making suitable probes or primers from the sequences provided

herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue.

The prostate cancer antigen polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified using techniques described herein or otherwise known in the art, such as, for example, by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural, synthetic or recombinant sources using techniques described herein or otherwise known in the art, such as, for example, antibodies of the invention raised against the polypeptides of the present invention in methods which are well known in the art.

The present invention provides a polynucleotide comprising, or alternatively consisting of, the nucleic acid sequence of SEQ ID NO:X, and/or a cDNA contained in ATCC Deposit No.:Z. The present invention also provides a polypeptide comprising, or alternatively, consisting of, the polypeptide sequence of SEQ ID NO:Y and/or a polypeptide encoded by the cDNA contained in ATCC Deposit No.:Z. Polynucleotides encoding a polypeptide comprising, or alternatively consisting of the polypeptide sequence of SEQ ID NO:Y and/or a polypeptide sequence encoded by the cDNA contained in ATCC Deposit No.:Z are also encompassed by the invention.

Polynucle tide and Polypeptide Variants

The present invention is directed to variants of the prostate cancer antigen polynucleotide sequence disclosed in SEQ ID NO:X, the complementary strand thereto, and/or the cDNA sequence contained in ATCC Deposit No.:Z.

The present invention also encompasses variants of the polypeptide sequence disclosed in SEQ ID NO:Y and/or encoded by the cDNA contained in ATCC Deposit No.:Z.

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

Thus, one aspect of the invention provides an isolated nucleic acid molecule comprising, or alternatively consisting of, a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a prostate cancer antigen polypeptide having an amino acid seqeunce as shown in the sequence listing and described in SEQ ID NO:X or encoded by the cDNA in ATCC Deposit No.:Z; (b) a nucleotide sequence encoding a mature prostate cancer antigen polypeptide having the amino acid sequence as shown in the sequence listing and described in SEQ ID NO:X or encoded by the cDNA in ATCC Deposit No.:Z; (c) a nucleotide sequence encoding a biologically active fragment of a prostate cancer antigen polypeptide having an amino acid sequence shown in the sequence listing and described in SEQ ID NO:X or encoded by the cDNA contained in ATCC Deposit No.:Z; (d) a nucleotide sequence encoding an antigenic fragment of a prostate cancer antigen polypeptide having an amino acid sequence shown in the sequence listing and described in SEQ ID NO:X or encoded by the cDNA conatined in ATCC Deposit No.:Z; (e) a nucleotide sequence encoding a prostate cancer antigen polypeptide comprising the complete amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:X or the cDNA contained in ATCC Deposit No.:Z; (f) a nucleotide sequence encoding a mature prostate cancer antigen polypeptide having the amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:X or the cDNA contained in ATCC Deposit No.:Z; (g) a nucleotide sequence encoding a biologically active fragment of a prostate cancer antigen polypeptide having an amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:X or the cDNA in ATCC Deposit No.:Z; (h) a nucleotide sequence encoding an antigenic fragment of a prostate cancer antigen polypeptide having an amino

acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:X or the cDNA in ATCC Deposit No.:Z; and (i) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above.

The present invention is also directed to nucleic acid molecules which comprise, or alternatively consist of, a nucleotide sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100%, identical to, for example, any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above, the nucleotide coding sequence in SEQ ID NO:X or the complementary strand thereto, the nucleotide coding sequence of the cDNA contained in a deposited cDNA clone or the complementary strand thereto, a nucleotide sequence encoding a polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO:X, a nucleotide sequence encoding the polypeptide of SEQ ID NO:Y, a nucleotide sequence encoding the polypeptide encoded by the cDNA contained in ATCC Deposit No.:Z, and/or polynucleotide fragments of any of these nucleic acid molecules (e.g., those fragments described herein). Polynucleotides which hybridize to the complement of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

Polypeptides encoded by these nucleic acid molecules are also encompassed by the invention. In another embodiment, the invention encompasses nucleic acid moleucles which comprise, or alternatively, consist of a polynucleotide which hybridizes under stringent hybridization conditions, or alternatively, under lower stringency conditions, to a polynucleotide in (a), (b), (c), (d), (e), (f), (g), or (h), above. Polynucleotides which hybridize to the complement of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

Another aspect of the invention provides an isolated nucleic acid molecule comprising, or alternatively consisting of, a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a prostate cancer antigen polypeptide having an amino acid sequence as shown in the sequence listing and described in Table 1; (b) a nucleotide sequence encoding a mature prostate cancer antigen polypeptide having the amino acid sequence as shown in the sequence listing and described in Table 1; (c) a nucleotide sequence encoding a biologically active fragment of a prostate cancer antigen polypeptide having an amino acid sequence shown in the sequence listing and

described in Table 1; (d) a nucleotide sequence encoding an antigenic fragment of a prostate cancer antigen polypeptide having an amino acid sequence shown in the sequence listing and described Table 1; (e) a nucleotide sequence encoding a prostate cancer antigen polypeptide comprising the complete amino acid sequence encoded by a human cDNA plasmid contained in a cDNA plasmid contained in the ATCC Deposit No.:Z and described in Table 1; (f) a nucleotide sequence encoding a mature prostate cancer antigen polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit No.:Z and described in Table 1; (g) a nucleotide sequence encoding a biologically active fragment of a prostate cancer antigen polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit No.:Z and described in Table 1; (h) a nucleotide sequence encoding an antigenic fragment of a prostate cancer antigen polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit No.:Z and described in Table 1; and (i) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above.

The present invention is also directed to nucleic acid molecules which comprise, or alternatively consist of, a nucleotide sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100%, identical to, for example, any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above.

The present invention is also directed to polypeptides which comprise, or alternatively consist of, an amino acid sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to, for example, the polypeptide sequence shown in SEQ ID NO:Y, the polypeptide sequence encoded by the cDNA contained in ATCC Deposit No.:Z, a polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO:X, and/or polypeptide fragments of any of these polypeptides (e.g., those fragments described herein). Polynucleotides which hybridize to the complement of the nucleic acid molecules encoding these polypeptides under stringent hybridization conditions or alternatively, under lower stringency conditions are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

By a nucleic acid having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the nucleic acid is identical to the reference sequence except that the nucleotide

sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a nucleic acid having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence referred to in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990)). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identiy are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the lenght of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score.

This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence referred to in Table 1 or a fragment thereof, or the amino acid sequence encoded by the cDNA contained in ATCC Deposit No.:Z or a fragment thereof, can be determined conventionally using

known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci.6:237- 245(1990)). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C- terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the

FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which less than 50, less than 40, less than 30, less than 20, less than 10, or 5-50, 5-25, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (for example, one skilled in the art may change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level and are included in the present invention. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the polypeptide of the present invention without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times

higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show a functional activity (e.g., substantial biological activity) of the polypeptides of the invention. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity.

The present application is directed to nucleic acid molecules at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to the nucleic acid sequences disclosed herein, (e.g., encoding a polypeptide having the amino acid sequence of an N and/or C terminal deletion), irrespective of whether they encode a polypeptide having functional activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having functional activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide

having functional activity include, inter alia, (1) isolating a gene or allelic or splice variants thereof in a cDNA library; (2) in situ hybridization (e.g., "FISH") to metaphase chromosomal spreads to provide precise chromosomal location of the gene, as described in Verma et al., Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York (1988); and (3) Northern Blot analysis for detecting mRNA expression in specific tissues.

Preferred, however, are nucleic acid molecules having sequences at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to the nucleic acid sequences disclosed herein, which do, in fact, encode a polypeptide having functional activity. By a polypeptide demonstrating a "functional activity" is meant, a polypeptide capable of displaying one or more known functional activities associated with a full-length (complete) protein. Such functional activities include, but are not limited to, biological activity, antigenicity [ability to bind (or compete with a polypeptide for binding) to an anti-polypeptide antibody], immunogenicity (ability to generate antibody which binds to a specific polypeptide of the invention), ability to form multimers with polypeptides of the invention, and ability to bind to a receptor or ligand for a polypeptide.

The functional activity of the polypeptides, and fragments, variants derivatives, and analogs thereof, can be assayed by various methods.

For example, in one embodiment where one is assaying for the ability to bind or compete with full-length polypeptide of the present invention for binding to anti-polypetide antibody, various immunoassays known in the art can be used, including but not limited to, competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitation reactions, immunodiffusion assays, in situ immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

In another embodiment, where a ligand is identified, or the ability of a polypeptide fragment, variant or derivative of the invention to multimerize is being evaluated, binding can be assayed, e.g., by means well-known in the art, such as, for example, reducing and non-reducing gel chromatography, protein affinity chromatography, and affinity blotting. See generally, Phizicky, E., et al., Microbiol. Rev. 59:94-123 (1995). In another embodiment, physiological correlates polypeptide of the present invention binding to its substrates (signal transduction) can be assayed.

In addition, assays described herein (see Examples) and otherwise known in the art may routinely be applied to measure the ability of polypeptides of the present invention and fragments, variants derivatives and analogs thereof to elicit polypeptide related biological activity (either in vitro or in vivo). Other methods will be known to the skilled artisan and are within the scope of the invention.

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to, for example, the nucleic acid sequence of the deposited clone, the nucleic acid sequence referred to in Table 1 (SEQ ID NO:X), or fragments thereof, will encode polypeptides "having functional activity." In fact, since degenerate variants of any of these nucleotide sequences all encode the same polypeptide, in many instances, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having functional activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid), as further described below.

For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different

species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly. Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as, for example, an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of a polypeptide having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course it is highly preferable for a polypeptide to have an amino acid sequence which comprises the amino acid sequence of a polypeptide of SEQ ID NO:Y, the amino acid sequence encoded by SEQ ID NO:X, and/or the amino acid sequence encoded by cDNA contained in ATCC Deposit No.:Z, in order of ever-increasing preference, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of SEQ ID NO:Y or fragments thereof (e.g., the mature form and/or other fragments described herein), the amino acid sequence encoded by SEQ ID NO:X or fragments thereof, and/or the amino acid sequence encoded by the cDNA contained in ATCC Deposit No.:Z or fragments thereof, is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

Polynucleotide and Polypeptide Fragments

The present invention is also directed to polynucleotide fragments of the polynucleotides (nucleic acids) of the invention. In the present invention, a "polynucleotide fragment" refers to a polynucleotide having a nucleic acid sequence which, for example: is a portion of the cDNA contained in a deposited clone, or a portion of a polynucleotide sequence encoding the amino acid sequence encoded by the cDNA in ATCC Deposit No.:Z; is a portion of the polynucleotide sequence in SEQ ID NO:X or the complementary strand thereto; is a polynucleotide sequence encoding a portion of a polypeptide encoded by SEQ ID NO:X; is a polynucleotide sequence encoding a portion of the polypeptide of SEQ ID NO:Y;

or is a portion of the polynucleotide sequence encoding the polypeptide of SEQ ID NO:Y. The nucleotide fragments of the invention are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt, at least about 50 nt, at least about 75 nt, or at least about 150 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the sequence contained in a deposited clone, or the nucleotide sequence shown in SEQ ID NO:X or the complementary stand thereto. In this context "about" includes the particularly recited value or a value larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. These nucleotide fragments have uses that include, but are not limited to, as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., at least 160, 170, 180, 190, 200, 250, 500, 600, 1000, or 2000 nucleotides in length) are also encompassed by the invention.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700,701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, 2001-2050, 2051-2100, 2101-2150, 2151-2200, 2201-2250, 2251-2300, 2301-2350, 2351-2400, 2401-2450, 2451-2500, 2501-2550, 2551-2600, 2601-2650, 2651-2700, 2701-2750, 2751-2800, 2801-2850, 2851-2900, 2901-2950, 2951-3000, 3001-3050, or 3051 to the end of SEQ ID NO:X, or the complementary strand thereto. In this context "about" includes the particularly recited range or a range larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has a functional activity (e.g., biological activity). More preferably, these polynucleotides can be used as probes or primers as discussed herein. Polynucleotides which hybridize to one or more of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, a sequence from

about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700,701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, 2001-2050, 2051-2100, 2101-2150, 2151-2200, 2201-2250, 2251-2300, 2301-2350, 2351-2400, 2401-2450, 2451-2500, 2501-2550, 2551-2600, 2601-2650, 2651-2700, 2701-2750, 2751-2800, 2801-2850, 2851-2900, 2901-2950, 2951-3000, 3001-3050, or 3051 to the end of the cDNA sequence contained in ATCC Deposit No.:Z, or the complementary strand thereto. In this context "about" includes the particularly recited range or a range larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has a functional activity (e.g., biological activity). More preferably, these polynucleotides can be used as probes or primers as discussed herein. Polynucleotides which hybridize to one or more of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

In the present invention, a "polypeptide fragment" refers to an amino acid sequence which is a portion of that contained in SEQ ID NO:Y, a portion of an amino acid sequence encoded by SEQ ID NO:X or the complement thereof, and/or a portion of an amino acid sequence encoded by the cDNA contained in ATCC Deposit No.:Z. Protein (polypeptide) fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 101-120, 121-140, 141-160, 161-180, 181-200, 201-220, 221-240, 241-260, 261-280, 281-300, 301-320, 321-340, 341-360, or 361 to the end of the coding region. Moreover, polypeptide fragments of the invention may be at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges or values, or ranges or values larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes. Polynucleotides encoding these polypeptide fragments are also encompassed by the invention.

Even if deletion of one or more amino acids from the N-terminus of a protein results in modification of loss of one or more biological functions of the protein, other functional activities (e.g., biological activities, ability to multimerize, ability to bind a ligand) may still be retained. For example, the ability of shortened muteins to induce and/or bind to antibodies which recognize the complete or mature forms of the polypeptides generally will be retained when less than the majority of the residues of the complete or mature polypeptide are removed from the N-terminus. Whether a particular polypeptide lacking N-terminal residues of a complete polypeptide retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art. It is not unlikely that a mutein with a large number of deleted N-terminal amino acid residues may retain some biological or immunogenic activities. In fact, peptides composed of as few as six amino acid residues may often evoke an immune response.

Accordingly, polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotides encoding these polypeptide fragments are also preferred.

The present invention further provides polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of a polypeptide disclosed herein (e.g., a polypeptide of SEQ ID NO:Y, a polypeptide encoded by the polynucleotide sequence contained in SEQ ID NO:X, and/or a polypeptide encoded by the cDNA contained in ATCC Deposit No.:Z). In particular, N-terminal deletions may be described by the general formula m-q, where q is a whole integer representing the total number of amino acid residues in a polypeptide of the invention (e.g., the polypeptide disclosed in SEQ ID NO:Y), and m is defined as any integer ranging from 2 to q-6. Polynucleotides encoding these polypeptides are also encompassed by the invention.

The present invention further provides polypeptides having one or more residues from the carboxy terminus of the amino acid sequence of a polypeptide disclosed herein (e.g., a

polypeptide of SEQ ID NO:Y, a polypeptide encoded by the polynucleotide sequence contained in SEQ ID NO:X, and/or a polypeptide encoded by the cDNA contained in ATCC Deposit No.:Z). In particular, C-terminal deletions may be described by the general formula 1-n, where n is any whole integer ranging from 6 to q-1, and where n corresponds to the position of amino acid residue in a polypeptide of the invention. Polynucleotides encoding these polypeptides are also encompassed by the invention.

In addition, any of the above described N- or C-terminal deletions can be combined to produce a N- and C-terminal deleted polypeptide. The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues m-n of a polypeptide encoded by SEQ ID NO:X (e.g., including, but not limited to, the preferred polypeptide disclosed as SEQ ID NO:Y), or the cDNA contained in ATCC Deposit No.:Z, and/or the complement thereof, where n and m are integers as described above. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Also as mentioned above, even if deletion of one or more amino acids from the C-terminus of a protein results in modification of loss of one or more biological functions of the protein, other functional activities (e.g., biological activities, ability to multimerize, ability to bind a ligand) may still be retained. For example the ability of the shortened mutein to induce and/or bind to antibodies which recognize the complete or mature forms of the polypeptide generally will be retained when less than the majority of the residues of the complete or mature polypeptide are removed from the C-terminus. Whether a particular polypeptide lacking C-terminal residues of a complete polypeptide retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art. It is not unlikely that a mutein with a large number of deleted C-terminal amino acid residues may retain some biological or immunogenic activities. In fact, peptides composed of as few as six amino acid residues may often evoke an immune response.

The present application is also directed to proteins containing polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the polypeptide sequence set forth. In preferred embodiments, the application is directed to proteins containing polypeptides at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to polypeptides having the amino acid sequence of the specific N- and C-terminal deletions. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Any polypeptide sequence encoded by the polynucleotide sequences set forth as SEQ ID NO:X (e.g., the polypeptide of SEQ ID NO::Y), or the cDNA contained in ATCC Deposit No.:Z may be analyzed to determine certain preferred regions of the polypeptide. For example, the amino acid sequence of a polypeptide encoded by a polynucleotide sequence of SEQ ID NO:X, or the cDNA contained in ATCC Deposit No.:Z may be analyzed using the default parameters of the DNASTAR computer algorithm (DNASTAR, Inc., 1228 S. Park St., Madison, WI 53715 USA; http://www.dnastar.com/).

Polypeptide regions that may be routinely obtained using the DNASTAR computer algorithm include, but are not limited to, Garnier-Robson alpha-regions, beta-regions, turn-regions, and coil-regions, Chou-Fasman alpha-regions, beta-regions, and turn-regions, Kyte-Doolittle hydrophilic regions and hydrophobic regions, Eisenberg alpha- and beta-amphipathic regions, Karplus-Schulz flexible regions, Emini surface-forming regions and Jameson-Wolf regions of high antigenic index. Among highly preferred polynucleotides of the invention in this regard are those that encode polypeptides comprising regions that combine several structural features, such as several (e.g., 1, 2, 3 or 4) of the features set out above.

Additionally, Kyte-Doolittle hydrophilic regions and hydrophobic regions, Emini surface-forming regions, and Jameson-Wolf regions of high antigenic index (i.e., containing four or more contiguous amino acids having an antigenic index of greater than or equal to 1.5, as identified using the default parameters of the Jameson-Wolf program) can routinely be used to determine polypeptide regions that exhibit a high degree of potential for antigenicity. Regions of high antigenicity are determined from data by DNASTAR analysis by choosing values which represent regions of the polypeptide which are likely to be exposed on the surface of the polypeptide in an environment in which antigen recognition may occur in the process of initiation of an immune response.

Preferred polypeptide fragments of the invention are fragments comprising, or alternatively, consisting of, an amino acid sequence that displays a functional activity (e.g., biological activity) of the polypeptide sequence of which the amino acid sequence is a fragment. By a polypeptide displaying a "functional acitivity" is meant a polypeptide capable of one or more known functional activities associated with a full length protein, such as, for example, biological activity, antigenicity, immunogenicity, and/or multimerization, as described herein.

Other preferred polypeptide fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

In preferred embodiments, polypeptides of the invention comprise, or alternatively consist of, one, two, three, four, five or more of the antigenic fragments of the polypeptide of SEQ ID NO:Y, or portions thereof. Polynucleotides encoding these polypeptides are also encompassed by the invention.

The present invention encompasses polypeptides comprising, or alternatively consisting of, an epitope of the polypeptide sequence shown in SEQ ID NO:Y, an epitope of a polypeptide sequence encoded by SEQ ID NO:X, an epitope of the polypeptide sequence encoded by the cDNA contained in ATCC Deposit No.:Z or an epitope of a polypeptide encoded by a polynucleotide that hybridizes to the complement of the sequence of SEQ ID NO:X, or the sequence of the cDNA contained in ATCC Deposit No.:Z under stringent hybridization conditions or alternatively, under lower stringency hybridization as defined supra. The present invention further encompasses polynucleotide sequences encoding an epitope of a polypeptide sequence of the invention (such as, for example, the sequence disclosed in SEQ ID NO:X, or a fragment thereof), polynucleotide sequences of the complementary strand of a polynucleotide sequence encoding an epitope of the invention, and polynucleotide sequences which hybridize to the complementary strand under stringent hybridization conditions or alternatively, under lower stringency hybridization conditions defined supra.

The term "epitopes," as used herein, refers to portions of a polypeptide having antigenic or immunogenic activity in an animal, preferably a mammal, and most preferably in a human. In a preferred embodiment, the present invention encompasses a polypeptide comprising an epitope, as well as the polynucleotide encoding this polypeptide. An "immunogenic epitope," as used herein, is defined as a portion of a protein that elicits an antibody response in an animal, as determined by any method known in the art, for example, by the methods for generating antibodies described infra. (See, for example, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983)). The term "antigenic epitope," as used herein, is defined as a portion of a protein to which an antibody can immunospecifically bind its antigen as determined by any method well known in the art, for example, by the

immunoassays described herein. Immunospecific binding excludes non-specific binding but does not necessarily exclude cross- reactivity with other antigens. Antigenic epitopes need not necessarily be immunogenic.

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least 4, at least 5, at least 6, at least 7, more preferably at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, and, most preferably, between about 15 to about 30 amino acids. Preferred polypeptides comprising immunogenic or antigenic epitopes are at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length. Additional non-exclusive preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as portions thereof. Antigenic epitopes are useful, for example, to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. Preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these antigenic epitopes. Antigenic epitopes can be used as the target molecules in immunoassays. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe et al., Science 219:660-666 (1983)).

Non-limiting examples of epitopes of polypeptides that can be used to generate antibodies of the invention include, a polypeptide comprising, or alternatively consisting of, at least one, two, three, four, five, six, seven, eight, nine, ten, or eleven or more of the portion(s) of SEQ ID NO:Y specified in the *Polynucleotides and Polypeptides of the Invention* section above. These polypeptide fragments have been determined to bear antigenic epitopes of the proteins of the invention by the analysis of the Jameson-Wolf antigenic index which is included in the DNAStar suite of computer programs. By "comprise" it is intended that a polypeptide contains at least one, two, three, four, five, six, seven, eight, nine, ten, or eleven or more of the portion(s) of SEQ ID NO:Y delineated above, but it may contain additional flanking residues on either the amino or carboxyl termini of the recited portion; i.e., contiguous sequence shown in SEQ ID NO:Y. The flanking sequence may, however, be sequence(s) from a heterologous polypeptide, such as from another protein described herein or from a heterologous polypeptide not described herein. In particular

embodiments, epitope portions of a polypeptide of the invention comprise one, two, three, or more of the portions of SEQ ID NO:Y described above

Similarly, immunogenic epitopes can be used, for example, to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle et al., J. Gen. Virol. 66:2347-2354 (1985). Preferred immunogenic epitopes include the immunogenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these immunogenic epitopes. The polypeptides comprising one or more immunogenic epitopes may be presented for eliciting an antibody response together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse), or, if the polypeptide is of sufficient length (at least about 25 amino acids), the polypeptide may be presented without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting).

Epitope-bearing polypeptides of the present invention may be used to induce antibodies according to methods well known in the art including, but not limited to, in vivo immunization, in vitro immunization, and phage display methods. See, e.g., Sutcliffe et al., supra; Wilson et al., supra, and Bittle et al., J. Gen. Virol., 66:2347-2354 (1985). If in vivo immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine residues may be coupled to a carrier using a linker such as maleimidobenzoyl- N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carriers using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier- coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg of peptide or carrier protein and Freund's adjuvant or any other adjuvant known for stimulating an immune response. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide

antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

As one of skill in the art will appreciate, and as discussed above, the polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to other polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, or any combination thereof and portions thereof) resulting in chimeric polypeptides. Such fusion proteins may facilitate purification and may increase half-life in vivo. This has been shown for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. See, e.g., EP 394,827; Traunecker et al., Nature, 331:84-86 (1988). Enhanced delivery of an antigen across the epithelial barrier to the immune system has been demonstrated for antigens (e.g., insulin) conjugated to an FcRn binding partner such as IgG or Fc fragments (see, e.g., PCT Publications WO 96/22024 and WO 99/04813). IgG Fusion proteins that have a disulfide-linked dimeric structure due to the IgG portion desulfide bonds have also been found to be more efficient in binding and neutralizing other molecules than monomeric polypeptides or fragments thereof alone. See, e.g., Fountoulakis et al., J. Biochem., 270:3958-3964 (1995). Nucleic acids encoding the above epitopes can also be recombined with a gene of interest as an epitope tag (e.g., the hemagglutinin ("HA") tag or flag tag) to aid in detection and purification of the expressed polypeptide. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., Proc. Natl. Acad. Sci. USA 88:8972- 897 (1991)). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with the recombinant vaccinia virus are loaded onto Ni2+ nitriloacetic acid-agarose column and histidine-tagged proteins can be selectively eluted with imidazole-containing buffers.

Additional fusion proteins of the invention may be generated through the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling"). DNA shuffling may be employed to modulate the activities

of polypeptides of the invention, such methods can be used to generate polypeptides with altered activity, as well as agonists and antagonists of the polypeptides. See, generally, U.S. Patent Nos. 5,605,793; 5,811,238; 5,830,721; 5,834,252; and 5,837,458, and Patten et al., Curr. Opinion Biotechnol. 8:724-33 (1997); Harayama, Trends Biotechnol. 16(2):76-82 (1998); Hansson, et al., J. Mol. Biol. 287:265-76 (1999); and Lorenzo and Blasco, Biotechniques 24(2):308- 13 (1998) (each of these patents and publications are hereby incorporated by reference in its entirety). In one embodiment, alteration of polynucleotides corresponding to SEQ ID NO:X and the polypeptides encoded by these polynucleotides may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments by homologous or site-specific recombination to generate variation in the polynucleotide sequence. In another embodiment, polynucleotides of the invention, or the encoded polypeptides, may be altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of a polynucleotide encoding a polypeptide of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, polypeptides of the present invention which are shown to be secreted can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

In certain preferred embodiments, proteins of the invention comprise fusion proteins wherein the polypeptides are N and/or C- terminal deletion mutants. In preferred

embodiments, the application is directed to nucleic acid molecules at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequences encoding polypeptides having the amino acid sequence of the specific N- and C-terminal deletions mutants. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

As one of skill in the art will appreciate, polypeptides of the present invention of the present invention and the epitope-bearing fragments thereof described above can be combined with heterologous polypeptide sequences. For example, the polypeptides of the present invention may be fused with heterologous polypeptide sequences, for example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM) or portions thereof (CH1, CH2, CH3, and any combination thereof, including both entire domains and portions thereof), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the

fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a polypeptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides of the invention may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable

promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, Streptomyces and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Preferred expression vectors for use in yeast systems include, but are not limited to pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, pPIC9K, and PAO815 (all available from Invitrogen, Carlbad, CA). Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In one embodiment, the yeast *Pichia pastoris* is used to express polypeptides of the invention in a eukaryotic system. *Pichia pastoris* is a methylotrophic yeast which can metabolize methanol as its sole carbon source. A main step in the methanol metabolization pathway is the oxidation of methanol to formaldehyde using O_2. This reaction is catalyzed by the enzyme alcohol oxidase. In order to metabolize methanol as its sole carbon source, *Pichia pastoris* must generate high levels of alcohol oxidase due, in part, to the relatively low affinity of alcohol oxidase for O_2. Consequently, in a growth medium depending on methanol as a main carbon source, the promoter region of one of the two alcohol oxidase genes (*AOX1*) is highly active. In the presence of methanol, alcohol oxidase produced from the *AOX1* gene comprises up to approximately 30% of the total soluble protein in *Pichia pastoris*. *See*, Ellis, S.B., *et al.*, *Mol. Cell. Biol.* 5:1111-21 (1985); Koutz, P.J, *et al.*, *Yeast*

5:167-77 (1989); Tschopp, J.F., *et al., Nucl. Acids Res.* 15:3859-76 (1987). Thus, a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, under the transcriptional regulation of all or part of the *AOX1* regulatory sequence is expressed at exceptionally high levels in *Pichia* yeast grown in the presence of methanol.

In one example, the plasmid vector pPIC9K is used to express DNA encoding a polypeptide of the invention, as set forth herein, in a *Pichea* yeast system essentially as described in "*Pichia* Protocols: Methods in Molecular Biology," D.R. Higgins and J. Cregg, eds. The Humana Press, Totowa, NJ, 1998. This expression vector allows expression and secretion of a polypeptide of the invention by virtue of the strong *AOX1* promoter linked to the *Pichia pastoris* alkaline phosphatase (PHO) secretory signal peptide (i.e., leader) located upstream of a multiple cloning site.

Many other yeast vectors could be used in place of pPIC9K, such as, pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalpha, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, and PAO815, as one skilled in the art would readily appreciate, as long as the proposed expression construct provides appropriately located signals for transcription, translation, secretion (if desired), and the like, including an in-frame AUG as required.

In another embodiment, high-level expression of a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, may be achieved by cloning the heterologous polynucleotide of the invention into an expression vector such as, for example, pGAPZ or pGAPZalpha, and growing the yeast culture in the absence of methanol.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994;

Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, Proteins: Structures and Molecular Principles, W.H. Freeman & Co., N.Y., and Hunkapiller et al., *Nature*, 310:105-111 (1984)). For example, a polypeptide corresponding to a fragment of a polypeptide can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the polypeptide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoro-amino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

Non-naturally occurring variants may be produced using art-known mutagenesis techniques, which include, but are not limited to oligonucleotide mediated mutagenesis, alanine scanning, PCR mutagenesis, site directed mutagenesis (*see, e.g.*, Carter *et al., Nucl. Acids Res. 13*:4331 (1986); and Zoller *et al., Nucl. Acids Res. 10*:6487 (1982)), cassette mutagenesis (*see, e.g.*, Wells *et al., Gene 34*:315 (1985)), restriction selection mutagenesis (*see, e.g.*, Wells *et al., Philos. Trans. R. Soc. London SerA 317*:415 (1986)).

The invention encompasses polypeptides of the present invention which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NaBH_4$; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

Also provided by the invention are chemically modified derivatives of the polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog). For example, the polyethylene glycol may have an average molecular weight of about 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10,000, 10,500, 11,000, 11,500, 12,000, 12,500, 13,000, 13,500, 14,000, 14,500, 15,000, 15,500, 16,000, 16,500, 17,000, 17,500, 18,000, 18,500, 19,000, 19,500, 20,000, 25,000, 30,000, 35,000, 40,000, 50,000, 55,000, 60,000, 65,000, 70,000, 75,000, 80,000, 85,000, 90,000, 95,000, or 100,000 kDa.

As noted above, the polyethylene glycol may have a branched structure. Branched polyethylene glycols are described, for example, in U.S. Patent No. 5,643,575; Morpurgo *et*

al., Appl. Biochem. Biotechnol. 56:59-72 (1996); Vorobjev *et al., Nucleosides Nucleotides 18*:2745-2750 (1999); and Caliceti *et al., Bioconjug. Chem. 10*:638-646 (1999), the disclosures of each of which are incorporated herein by reference.

The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., Exp. Hematol. 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

As suggested above, polyethylene glycol may be attached to proteins via linkage to any of a number of amino acid residues. For example, polyethylene glycol can be linked to a proteins via covalent bonds to lysine, histidine, aspartic acid, glutamic acid, or cysteine residues. One or more reaction chemistries may be employed to attach polyethylene glycol to specific amino acid residues (e.g., lysine, histidine, aspartic acid, glutamic acid, or cysteine) of the protein or to more than one type of amino acid residue (e.g., lysine, histidine, aspartic acid, glutamic acid, cysteine and combinations thereof) of the protein.

One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus

modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

As indicated above, pegylation of the proteins of the invention may be accomplished by any number of means. For example, polyethylene glycol may be attached to the protein either directly or by an intervening linker. Linkerless systems for attaching polyethylene glycol to proteins are described in Delgado *et al., Crit. Rev. Thera. Drug Carrier Sys. 9*:249-304 (1992); Francis *et al., Intern. J. of Hematol. 68*:1-18 (1998); U.S. Patent No. 4,002,531; U.S. Patent No. 5,349,052; WO 95/06058; and WO 98/32466, the disclosures of each of which are incorporated herein by reference.

One system for attaching polyethylene glycol directly to amino acid residues of proteins without an intervening linker employs tresylated MPEG, which is produced by the modification of monmethoxy polyethylene glycol (MPEG) using tresylchloride ($ClSO_2CH_2CF_3$). Upon reaction of protein with tresylated MPEG, polyethylene glycol is directly attached to amine groups of the protein. Thus, the invention includes protein-polyethylene glycol conjugates produced by reacting proteins of the invention with a polyethylene glycol molecule having a 2,2,2-trifluoreothane sulphonyl group.

Polyethylene glycol can also be attached to proteins using a number of different intervening linkers. For example, U.S. Patent No. 5,612,460, the entire disclosure of which is incorporated herein by reference, discloses urethane linkers for connecting polyethylene glycol to proteins. Protein-polyethylene glycol conjugates wherein the polyethylene glycol is attached to the protein by a linker can also be produced by reaction of proteins with compounds such as MPEG-succinimidylsuccinate, MPEG activated with 1,1'-carbonyldiimidazole, MPEG-2,4,5-trichloropenylcarbonate, MPEG-p-nitrophenolcarbonate, and various MPEG-succinate derivatives. A number additional polyethylene glycol derivatives and reaction chemistries for attaching polyethylene glycol to proteins are described in WO 98/32466, the entire disclosure of which is incorporated herein by reference. Pegylated protein products produced using the reaction chemistries set out herein are included within the scope of the invention.

The number of polyethylene glycol moieties attached to each protein of the invention (*i.e.*, the degree of substitution) may also vary. For example, the pegylated proteins of the invention may be linked, on average, to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20, or more polyethylene glycol molecules. Similarly, the average degree of substitution within ranges such as 1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9, 8-10, 9-11, 10-12, 11-13, 12-14, 13-15, 14-16, 15-17, 16-18, 17-19, or 18-20 polyethylene glycol moieties per protein molecule. Methods for determining the degree of substitution are discussed, for example, in Delgado *et al.*, *Crit. Rev. Thera. Drug Carrier Sys.* 9:249-304 (1992).

The polypeptides of the invention may be in monomers or multimers (i.e., dimers, trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the polypeptides of the invention, their preparation, and compositions (preferably, Therapeutics) containing them. In specific embodiments, the polypeptides of the invention are monomers, dimers, trimers or tetramers. In additional embodiments, the multimers of the invention are at least dimers, at least trimers, or at least tetramers.

Multimers encompassed by the invention may be homomers or heteromers. As used herein, the term homomer, refers to a multimer containing only polypeptides corresponding to the amino acid sequence of SEQ ID NO:Y, an amino acid sequence encoded by SEQ ID NO:X and/or an amino acid sequence encoded by the cDNA contained in ATCC Deposit No.:Z (including fragments, variants, splice variants, and fusion proteins, corresponding to these as described herein). These homomers may contain polypeptides having identical or different amino acid sequences. In a specific embodiment, a homomer of the invention is a multimer containing only polypeptides having an identical amino acid sequence. In another specific embodiment, a homomer of the invention is a multimer containing polypeptides having different amino acid sequences. In specific embodiments, the multimer of the invention is a homodimer (e.g., containing polypeptides having identical or different amino acid sequences) or a homotrimer (e.g., containing polypeptides having identical and/or different amino acid sequences). In additional embodiments, the homomeric multimer of the invention is at least a homodimer, at least a homotrimer, or at least a homotetramer.

As used herein, the term heteromer refers to a multimer containing one or more heterologous polypeptides (i.e., polypeptides of different proteins) in addition to the polypeptides of the invention. In a specific embodiment, the multimer of the invention is a

heterodimer, a heterotrimer, or a heterotetramer. In additional embodiments, the heteromeric multimer of the invention is at least a heterodimer, at least a heterotrimer, or at least a heterotetramer.

Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked, by for example, liposome formation. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the invention (including antibodies to the heterologous polypeptide sequence in a fusion protein of the invention) in solution. In other embodiments, multimers of the invention are formed by covalent associations with and/or between the polypeptides of the invention. Such covalent associations may involve one or more amino acid residues contained in the polypeptide sequence (e.g., that recited in SEQ ID NO:Y, or contained in a polypeptide encoded by SEQ ID NO:X, and/or the deposited cDNA clone). In one instance, the covalent associations are cross-linking between cysteine residues located within the polypeptide sequences which interact in the native (i.e., naturally occurring) polypeptide. In another instance, the covalent associations are the consequence of chemical or recombinant manipulation. Alternatively, such covalent associations may involve one or more amino acid residues contained in the heterologous polypeptide sequence in a fusion protein. In one example, covalent associations are between the heterologous sequence contained in a fusion protein of the invention (see, e.g., US Patent Number 5,478,925). In a specific example, the covalent associations are between the heterologous sequence contained in a Fc fusion protein of the invention (as described herein). In another specific example, covalent associations of fusion proteins of the invention are between heterologous polypeptide sequence from another protein that is capable of forming covalently associated multimers, such as for example, oseteoprotegerin (see, e.g., International Publication NO: WO 98/49305, the contents of which are herein incorporated by reference in its entirety). In another embodiment, two or more polypeptides of the invention are joined through peptide linkers. Examples include those peptide linkers described in U.S. Pat. No. 5,073,627 (hereby incorporated by reference). Proteins comprising multiple polypeptides of

the invention separated by peptide linkers may be produced using conventional recombinant DNA technology.

Another method for preparing multimer polypeptides of the invention involves use of polypeptides of the invention fused to a leucine zipper or isoleucine zipper polypeptide sequence. Leucine zipper and isoleucine zipper domains are polypeptides that promote multimerization of the proteins in which they are found. Leucine zippers were originally identified in several DNA-binding proteins (Landschulz et al., Science 240:1759, (1988)), and have since been found in a variety of different proteins. Among the known leucine zippers are naturally occurring peptides and derivatives thereof that dimerize or trimerize. Examples of leucine zipper domains suitable for producing soluble multimeric proteins of the invention are those described in PCT application WO 94/10308, hereby incorporated by reference. Recombinant fusion proteins comprising a polypeptide of the invention fused to a polypeptide sequence that dimerizes or trimerizes in solution are expressed in suitable host cells, and the resulting soluble multimeric fusion protein is recovered from the culture supernatant using techniques known in the art.

Trimeric polypeptides of the invention may offer the advantage of enhanced biological activity. Preferred leucine zipper moieties and isoleucine moieties are those that preferentially form trimers. One example is a leucine zipper derived from lung surfactant protein D (SPD), as described in Hoppe et al. (FEBS Letters 344:191, (1994)) and in U.S. patent application Ser. No. 08/446,922, hereby incorporated by reference. Other peptides derived from naturally occurring trimeric proteins may be employed in preparing trimeric polypeptides of the invention.

In another example, proteins of the invention are associated by interactions between Flag® polypeptide sequence contained in fusion proteins of the invention containing Flag® polypeptide seuqence. In a further embodiment, associations proteins of the invention are associated by interactions between heterologous polypeptide sequence contained in Flag® fusion proteins of the invention and anti-Flag® antibody.

The multimers of the invention may be generated using chemical techniques known in the art. For example, polypeptides desired to be contained in the multimers of the invention may be chemically cross-linked using linker molecules and linker molecule length optimization techniques known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, multimers of the invention

may be generated using techniques known in the art to form one or more inter-molecule cross-links between the cysteine residues located within the sequence of the polypeptides desired to be contained in the multimer (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Further, polypeptides of the invention may be routinely modified by the addition of cysteine or biotin to the C-terminus or N-terminus of the polypeptide and techniques known in the art may be applied to generate multimers containing one or more of these modified polypeptides (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, techniques known in the art may be applied to generate liposomes containing the polypeptide components desired to be contained in the multimer of the invention (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

Alternatively, multimers of the invention may be generated using genetic engineering techniques known in the art. In one embodiment, polypeptides contained in multimers of the invention are produced recombinantly using fusion protein technology described herein or otherwise known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In a specific embodiment, polynucleotides coding for a homodimer of the invention are generated by ligating a polynucleotide sequence encoding a polypeptide of the invention to a sequence encoding a linker polypeptide and then further to a synthetic polynucleotide encoding the translated product of the polypeptide in the reverse orientation from the original C-terminus to the N-terminus (lacking the leader sequence) (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In another embodiment, recombinant techniques described herein or otherwise known in the art are applied to generate recombinant polypeptides of the invention which contain a transmembrane domain (or hyrophobic or signal peptide) and which can be incorporated by membrane reconstitution techniques into liposomes (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

Antibodies

Further polypeptides of the invention relate to antibodies and T-cell antigen receptors (TCR) which immunospecifically bind a polypeptide, polypeptide fragment, or variant of SEQ ID NO:Y, and/or an epitope, of the present invention (as determined by immunoassays well known in the art for assaying specific antibody-antigen binding). Antibodies of the

invention include, but are not limited to, polyclonal, monoclonal, multispecific, human, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies (including, e.g., anti-Id antibodies to antibodies of the invention), and epitope-binding fragments of any of the above. The term "antibody," as used herein, refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that immunospecifically binds an antigen. The immunoglobulin molecules of the invention can be of any type (e.g., IgG, IgE, IgM, IgD, IgA and IgY), class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass of immunoglobulin molecule. In preferred embodiments, the immunoglobulin molecules of the invention are IgG1. In other preferred embodiments, the immunoglobulin molecules of the invention are IgG4.

Most preferably the antibodies are human antigen-binding antibody fragments of the present invention and include, but are not limited to, Fab, Fab' and F(ab')2, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a VL or VH domain. Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entirety or a portion of the following: hinge region, CH1, CH2, and CH3 domains. Also included in the invention are antigen-binding fragments also comprising any combination of variable region(s) with a hinge region, CH1, CH2, and CH3 domains. The antibodies of the invention may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine (e.g., mouse and rat), donkey, ship rabbit, goat, guinea pig, camel, horse, or chicken. As used herein, "human" antibodies include antibodies having the amino acid sequence of a human immunoglobulin and include antibodies isolated from human immunoglobulin libraries or from animals transgenic for one or more human immunoglobulin and that do not express endogenous immunoglobulins, as described infra and, for example in, U.S. Patent No. 5,939,598 by Kucherlapati et al.

The antibodies of the present invention may be monospecific, bispecific, trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for a heterologous epitope, such as a heterologous polypeptide or solid support material. See, e.g., PCT publications WO 93/17715; WO 92/08802; WO 91/00360;

WO 92/05793; Tutt, et al., J. Immunol. 147:60-69 (1991); U.S. Patent Nos. 4,474,893; 4,714,681; 4,925,648; 5,573,920; 5,601,819; Kostelny et al., J. Immunol. 148:1547-1553 (1992).

Antibodies of the present invention may be described or specified in terms of the epitope(s) or portion(s) of a polypeptide of the present invention which they recognize or specifically bind. The epitope(s) or polypeptide portion(s) may be specified as described herein, e.g., by N-terminal and C-terminal positions, or by size in contiguous amino acid residues, or specified in the section entitled *Polynucleotides and Polypeptides of the Invention* for each SEQ ID NO:Y. Preferred epitopes of the invention include those shown in above, also preferred are polynucleotides that encode these eptiopes. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded. Therefore, the present invention includes antibodies that specifically bind polypeptides of the present invention, and allows for the exclusion of the same.

Antibodies of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies that do not bind any other analog, ortholog, or homolog of a polypeptide of the present invention are included. Antibodies that bind polypeptides with at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, and at least 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In specific embodiments, antibodies of the present invention cross-react with murine, rat and/or rabbit homologs of human proteins and the corresponding epitopes thereof. Antibodies that do not bind polypeptides with less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, and less than 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In a specific embodiment, the above-described cross-reactivity is with respect to any single specific antigenic or immunogenic polypeptide, or combination(s) of 2, 3, 4, 5, or more of the specific antigenic and/or immunogenic polypeptides disclosed herein. Further included in the present invention are antibodies which bind polypeptides encoded by polynucleotides which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein). Antibodies of the present invention may also be described or specified in terms of their binding affinity to a polypeptide of the invention. Preferred

binding affinities include those with a dissociation constant or Kd less than 5 X 10^{-2} M, 10^{-2} M, 5 X 10^{-3} M, 10^{-3} M, 5 X 10^{-4} M, 10^{-4} M, 5 X 10^{-5} M, 10^{-5} M, 5 X 10^{-6} M, 10^{-6}M, 5 X 10^{-7} M, 10^{7} M, 5 X 10^{-8} M, 10^{-8} M, 5 X 10^{-9} M, 10^{-9} M, 5 X 10^{-10} M, 10^{-10} M, 5 X 10^{-11} M, 10^{-11} M, 5 X 10^{-12} M, $^{10-12}$ M, 5 X 10^{-13} M, 10^{-13} M, 5 X 10^{-14} M, 10^{-14} M, 5 X 10^{-15} M, or $^{10-15}$ M.

The invention also provides antibodies that competitively inhibit binding of an antibody to an epitope of the invention as determined by any method known in the art for determining competitive binding, for example, the immunoassays described herein. In preferred embodiments, the antibody competitively inhibits binding to the epitope by at least 95%, at least 90%, at least 85 %, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50%.

Antibodies of the present invention may act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes antibodies which disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. Preferrably, antibodies of the present invention bind an antigenic epitope disclosed herein, or a portion thereof. The invention features both receptor-specific antibodies and ligand-specific antibodies. The invention also features receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. For example, receptor activation can be determined by detecting the phosphorylation (e.g., tyrosine or serine/threonine) of the receptor or its substrate by immunoprecipitation followed by western blot analysis (for example, as described supra). In specific embodiments, antibodies are provided that inhibit ligand activity or receptor activity by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50% of the activity in absence of the antibody.

The invention also features receptor-specific antibodies which both prevent ligand binding and receptor activation as well as antibodies that recognize the receptor-ligand complex, and, preferably, do not specifically recognize the unbound receptor or the unbound ligand. Likewise, included in the invention are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor activation, but do not prevent the ligand from binding the receptor. Further included in the invention are antibodies which activate the receptor. These antibodies may act as receptor agonists, i.e., potentiate or activate either all or a subset of the

biological activities of the ligand-mediated receptor activation, for example, by inducing dimerization of the receptor. The antibodies may be specified as agonists, antagonists or inverse agonists for biological activities comprising the specific biological activities of the peptides of the invention disclosed herein. The above antibody agonists can be made using methods known in the art. See, e.g., PCT publication WO 96/40281; U.S. Patent No. 5,811,097; Deng et al., Blood 92(6):1981-1988 (1998); Chen et al., Cancer Res. 58(16):3668-3678 (1998); Harrop et al., J. Immunol. 161(4):1786-1794 (1998); Zhu et al., Cancer Res. 58(15):3209-3214 (1998); Yoon et al., J. Immunol. 160(7):3170-3179 (1998); Prat et al., J. Cell. Sci. 111(Pt2):237-247 (1998); Pitard et al., J. Immunol. Methods 205(2):177-190 (1997); Liautard et al., Cytokine 9(4):233-241 (1997); Carlson et al., J. Biol. Chem. 272(17):11295-11301 (1997); Taryman et al., Neuron 14(4):755-762 (1995); Muller et al., Structure 6(9):1153-1167 (1998); Bartunek et al., Cytokine 8(1):14-20 (1996) (which are all incorporated by reference herein in their entireties).

Antibodies of the present invention may be used, for example, but not limited to, to purify, detect, and target the polypeptides of the present invention, including both in vitro and in vivo diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of the polypeptides of the present invention in biological samples. See, e.g., Harlow et al., Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988) (incorporated by reference herein in its entirety).

As discussed in more detail below, the antibodies of the present invention may be used either alone or in combination with other compositions. The antibodies may further be recombinantly fused to a heterologous polypeptide at the N- or C-terminus or chemically conjugated (including covalently and non-covalently conjugations) to polypeptides or other compositions. For example, antibodies of the present invention may be recombinantly fused or conjugated to molecules useful as labels in detection assays and effector molecules such as heterologous polypeptides, drugs, radionuclides, or toxins. See, e.g., PCT publications WO 92/08495; WO 91/14438; WO 89/12624; U.S. Patent No. 5,314,995; and EP 396,387.

The antibodies of the invention include derivatives that are modified, i.e, by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from generating an anti-idiotypic response. For example, but not by way of limitation, the antibody derivatives include antibodies that have been modified,

e.g., by glycosylation, acetylation, pegylation, phosphylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, etc. Additionally, the derivative may contain one or more non-classical amino acids.

The antibodies of the present invention may be generated by any suitable method known in the art. Polyclonal antibodies to an antigen-of- interest can be produced by various procedures well known in the art. For example, a polypeptide of the invention can be administered to various host animals including, but not limited to, rabbits, mice, rats, etc. to induce the production of sera containing polyclonal antibodies specific for the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. Such adjuvants are also well known in the art.

Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow et al., Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: Monoclonal Antibodies and T-Cell Hybridomas 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). The term "monoclonal antibody" as used herein is not limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art and are discussed in detail in the Examples. In a non-limiting example, mice can be immunized with a polypeptide of the invention or a cell expressing such peptide. Once an immune response is detected, e.g., antibodies specific

for the antigen are detected in the mouse serum, the mouse spleen is harvested and splenocytes isolated. The splenocytes are then fused by well known techniques to any suitable myeloma cells, for example cells from cell line SP20 available from the ATCC. Hybridomas are selected and cloned by limited dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding a polypeptide of the invention. Ascites fluid, which generally contains high levels of antibodies, can be generated by immunizing mice with positive hybridoma clones.

Accordingly, the present invention provides methods of generating monoclonal antibodies as well as antibodies produced by the method comprising culturing a hybridoma cell secreting an antibody of the invention wherein, preferably, the hybridoma is generated by fusing splenocytes isolated from a mouse immunized with an antigen of the invention with myeloma cells and then screening the hybridomas resulting from the fusion for hybridoma clones that secrete an antibody able to bind a polypeptide of the invention.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, Fab and F(ab')2 fragments of the invention may be produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). F(ab')2 fragments contain the variable region, the light chain constant region and the CH1 domain of the heavy chain.

For example, the antibodies of the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular embodiment, such phage can be utilized to display antigen binding domains expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, e.g., using labeled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186

(1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187 9-18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT application No. PCT/GB91/01134; PCT publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')2 fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., BioTechniques 12(6):864-869 (1992); and Sawai et al., AJRI 34:26-34 (1995); and Better et al., Science 240:1041-1043 (1988) (said references incorporated by reference in their entireties).

Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston et al., Methods in Enzymology 203:46-88 (1991); Shu et al., PNAS 90:7995-7999 (1993); and Skerra et al., Science 240:1038-1040 (1988). For some uses, including in vivo use of antibodies in humans and in vitro detection assays, it may be preferable to use chimeric, humanized, or human antibodies. A chimeric antibody is a molecule in which different portions of the antibody are derived from different animal species, such as antibodies having a variable region derived from a murine monoclonal antibody and a human immunoglobulin constant region. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Gillies et al., (1989) J. Immunol. Methods 125:191-202; U.S. Patent Nos. 5,807,715; 4,816,567; and 4,816397, which are incorporated herein by reference in their entirety. Humanized antibodies are antibody molecules from non-human species antibody that binds the desired antigen having one or more complementarity determining regions (CDRs) from the non-human species and a framework regions from a human immunoglobulin molecule. Often, framework residues in the human framework regions will be substituted with the corresponding residue from the

CDR donor antibody to alter, preferably improve, antigen binding. These framework substitutions are identified by methods well known in the art, e.g., by modeling of the interactions of the CDR and framework residues to identify framework residues important for antigen binding and sequence comparison to identify unusual framework residues at particular positions. (See, e.g., Queen et al., U.S. Patent No. 5,585,089; Riechmann et al., Nature 332:323 (1988), which are incorporated herein by reference in their entireties.) Antibodies can be humanized using a variety of techniques known in the art including, for example, CDR-grafting (EP 239,400; PCT publication WO 91/09967; U.S. Patent Nos. 5,225,539; 5,530,101; and 5,585,089), veneering or resurfacing (EP 592,106; EP 519,596; Padlan, Molecular Immunology 28(4/5):489-498 (1991); Studnicka et al., Protein Engineering 7(6):805-814 (1994); Roguska. et al., PNAS 91:969-973 (1994)), and chain shuffling (U.S. Patent No. 5,565,332).

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Human antibodies can be made by a variety of methods known in the art including phage display methods described above using antibody libraries derived from human immunoglobulin sequences. See also, U.S. Patent Nos. 4,444,887 and 4,716,111; and PCT publications WO 98/46645, WO 98/50433, WO 98/24893, WO 98/16654, WO 96/34096, WO 96/33735, and WO 91/10741; each of which is incorporated herein by reference in its entirety.

Human antibodies can also be produced using transgenic mice which are incapable of expressing functional endogenous immunoglobulins, but which can express human immunoglobulin genes. For example, the human heavy and light chain immunoglobulin gene complexes may be introduced randomly or by homologous recombination into mouse embryonic stem cells. Alternatively, the human variable region, constant region, and diversity region may be introduced into mouse embryonic stem cells in addition to the human heavy and light chain genes. The mouse heavy and light chain immunoglobulin genes may be rendered non-functional separately or simultaneously with the introduction of human immunoglobulin loci by homologous recombination. In particular, homozygous deletion of the JH region prevents endogenous antibody production. The modified embryonic stem cells are expanded and microinjected into blastocysts to produce chimeric mice. The chimeric mice are then bred to produce homozygous offspring which express human antibodies. The transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a

portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained from the immunized, transgenic mice using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar, Int. Rev. Immunol. 13:65-93 (1995). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., PCT publications WO 98/24893; WO 92/01047; WO 96/34096; WO 96/33735; European Patent No. 0 598 877; U.S. Patent Nos. 5,413,923; 5,625,126; 5,633,425; 5,569,825; 5,661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771; and 5,939,598, which are incorporated by reference herein in their entirety. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al., Bio/technology 12:899-903 (1988)).

Further, antibodies to the polypeptides of the invention can, in turn, be utilized to generate anti-idiotype antibodies that "mimic" polypeptides of the invention using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, FASEB J. 7(5):437-444; (1989) and Nissinoff, J. Immunol. 147(8):2429-2438 (1991)). For example, antibodies which bind to and competitively inhibit polypeptide multimerization and/or binding of a polypeptide of the invention to a ligand can be used to generate anti-idiotypes that "mimic" the polypeptide multimerization and/or binding domain and, as a consequence, bind to and neutralize polypeptide and/or its ligand. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypes can be used in therapeutic regimens to neutralize polypeptide ligand. For example, such anti-idiotypic antibodies can be used to bind a polypeptide of the invention and/or to bind its ligands/receptors, and thereby block its biological activity.

Polynucleotides Encoding Antibodies

The invention further provides polynucleotides comprising a nucleotide sequence encoding an antibody of the invention and fragments thereof. The invention also encompasses polynucleotides that hybridize under stringent or alternatively, under lower stringency hybridization conditions, e.g., as defined supra, to polynucleotides that encode an antibody, preferably, that specifically binds to a polypeptide of the invention, preferably, an antibody that binds to a polypeptide having the amino acid sequence of SEQ ID NO:Y, and/or to a polypeptide encoded by the cDNA contained in ATCC Deposit No.:Z.

The polynucleotides may be obtained, and the nucleotide sequence of the polynucleotides determined, by any method known in the art. For example, if the nucleotide sequence of the antibody is known, a polynucleotide encoding the antibody may be assembled from chemically synthesized oligonucleotides (e.g., as described in Kutmeier et al., BioTechniques 17:242 (1994)), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence encoding the antibody, annealing and ligating of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

Alternatively, a polynucleotide encoding an antibody may be generated from nucleic acid from a suitable source. If a clone containing a nucleic acid encoding a particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the immunoglobulin may be chemically synthesized or obtained from a suitable source (e.g., an antibody cDNA library, or a cDNA library generated from, or nucleic acid, preferably poly A+ RNA, isolated from, any tissue or cells expressing the antibody, such as hybridoma cells selected to express an antibody of the invention) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence to identify, e.g., a cDNA clone from a cDNA library that encodes the antibody. Amplified nucleic acids generated by PCR may then be cloned into replicable cloning vectors using any method well known in the art.

Once the nucleotide sequence and corresponding amino acid sequence of the antibody is determined, the nucleotide sequence of the antibody may be manipulated using methods well known in the art for the manipulation of nucleotide sequences, e.g., recombinant DNA techniques, site directed mutagenesis, PCR, etc. (see, for example, the techniques described

in Sambrook et al., 1990, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY and Ausubel et al., eds., 1998, Current Protocols in Molecular Biology, John Wiley & Sons, NY, which are both incorporated by reference herein in their entireties), to generate antibodies having a different amino acid sequence, for example to create amino acid substitutions, deletions, and/or insertions.

In a specific embodiment, the amino acid sequence of the heavy and/or light chain variable domains may be inspected to identify the sequences of the complementarity determining regions (CDRs) by methods that are well know in the art, e.g., by comparison to known amino acid sequences of other heavy and light chain variable regions to determine the regions of sequence hypervariability. Using routine recombinant DNA techniques, one or more of the CDRs may be inserted within framework regions, e.g., into human framework regions to humanize a non-human antibody, as described supra. The framework regions may be naturally occurring or consensus framework regions, and preferably human framework regions (see, e.g., Chothia et al., J. Mol. Biol. 278: 457-479 (1998) for a listing of human framework regions). Preferably, the polynucleotide generated by the combination of the framework regions and CDRs encodes an antibody that specifically binds a polypeptide of the invention. Preferably, as discussed supra, one or more amino acid substitutions may be made within the framework regions, and, preferably, the amino acid substitutions improve binding of the antibody to its antigen. Additionally, such methods may be used to make amino acid substitutions or deletions of one or more variable region cysteine residues participating in an intrachain disulfide bond to generate antibody molecules lacking one or more intrachain disulfide bonds. Other alterations to the polynucleotide are encompassed by the present invention and within the skill of the art.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., Proc. Natl. Acad. Sci. 81:851-855 (1984); Neuberger et al., Nature 312:604-608 (1984); Takeda et al., Nature 314:452-454 (1985)) by splicing genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described supra, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region, e.g., humanized antibodies.